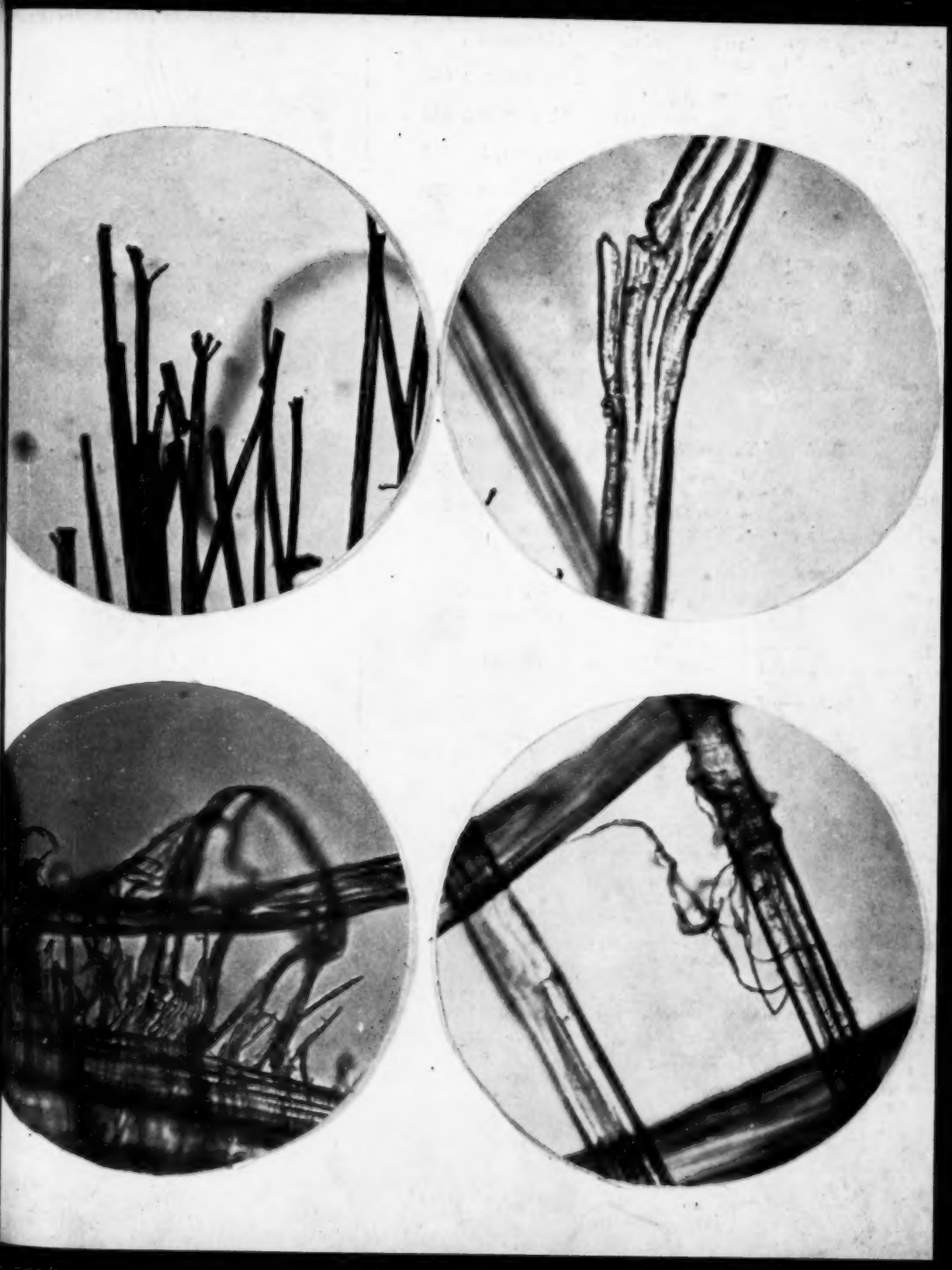
JOURNAL OF APPLIED PHYSICS



JOURNAL OF APPLIED PHYSICS

Vol. 17, No. 6 JUNE, 1946

In This Issue

Page

Contributed Original Research

Study of Distortion in Electron Microscope Projection Lenses JAMES HILLIER 411

Parafocusing Properties of Microcrystalline Powder Layers in X-Ray Diffraction Applied to the Design of X-Ray Goniometers J. C. M. BRENTANO 420

Study of Oxide Cathodes by X-Ray Diffraction Methods. Part I. Methods, Conversion Studies, and Thermal Expansion Coefficients A. EISENSTEIN 434

Methods for Betatron or Synchrotron Beam Removal EUGENE C. CRITTENDEN, JR. AND WILLIAM E. PARKINS 444

Graphical Method for Determining Particle Trajectories WILLIAM E. PARKINS AND EUGENE C. CRITTENDEN, JR. 447

Analyses of Mixtures of Light Gases by Infra-Red Absorption NORMAN D. COGGESHALL AND ELEANOR L. SAIER 450

Methods of Minimizing Lead Loss in Emissivity and Resistivity Determinations L. MALTER AND D. B. LANGMUIR 456

Theory and Application of the Parallel Plate Plastometer G. J. DIENES AND H. F. KLEMM 458

General Relations of Flow in Solids and Their Application to the Plastic Behavior of Tire Cords W. JAMES LYONS 472

Fatigue Failure of Rayon Tire Cord R. C. WALLER AND W. E. ROSEVEARE 482

Dynamic Characteristics of Rubber Supports from Vibration Table Data ERNEST G. CHILTON 492

Absorption of Microwaves by Gases W. D. HERSHBERGER 495

Pressure-Volume Relation for Cylindrical Tubes with Elastomeric Walls: The Human Aorta ALLEN L. KING 501

Approximations to Planckian Distributions PARRY MOON AND DOMINA EBERLE SPENCER 506

Approximate Calculation of the Range of a Winged Jet Projectile M. ZBIGMEW KRZYWOBLOCKI 515

Design of Broad Band I.F. Amplifiers RICHARD F. BAUM 519

Relation between Nodal Positions and Standing Wave Ratio in a Composite Transmission System EUGENE FEENBERG 530

Erratum: Small Prism Infra-Red Spectroscopy R. BOWLING BARNES, ROBERT S. MCDONALD, AND VAN ZANDT WILLIAMS 532

In Every Issue

Calendar of Meetings 514
Here and There 533
New Books 535
New Booklets 540
Index to Advertisers xxx

BOARD OF EDITORS

ELMER HUTCHISSON, *Editor*

Associate Editors

R. BOWLING BARNES
OTTO BEECK
W. F. BUSSE
S. H. CALDWELL
CLARK GOODMAN
SELIG HECHT
W. JAMES LYONS
PAUL E. SABINE
FREDERICK SEITZ
J. R. TOWNSEND

Publication of the American Institute of Physics

ADMINISTRATIVE STAFF

HENRY A. BARTON, *Director*

MARGARET S. GRIFFIN *Publications Manager*

THEODORE VORBURGER, *Business Manager*

JOHN T. TATE, *Adviser on Publications*

GOVERNING BOARD

PAUL E. KLOPSTEG, *Chairman*

WALLACE WATERFALL, *Secretary*

GEORGE B. PEGRAM, *Treasurer*

R. B. DOW
L. A. DUBRIDGE
W. F. FAIR, JR.
FLOYD A. FIRESTONE
HARVEY FLETCHER
R. C. GIBBS
K. S. GIBSON
W. H. MARKWOOD, JR.
A. H. PFUND
JOHN T. TATE
L. W. TAYLOR
A. G. WORTHING

The JOURNAL OF APPLIED PHYSICS, published monthly at Prince and Lemon Sts., Lancaster, Pa., is devoted to physics in its role as the science basic to other natural sciences and to the arts and industries. Previous to 1937 this publication was known as PHYSICS. The JOURNAL OF APPLIED PHYSICS publishes editorials, news of physicists, and reviews, as well as technical papers of applied physics.

Manuscripts should be submitted to Elmer Hutchisson, Editor, Case School of Applied Science, Cleveland 6, Ohio.

Subscriptions and orders for back numbers should be addressed to Prince and Lemon Streets, Lancaster, Pa. or to the American Institute of Physics, 57 East 55 Street, New York 22, N. Y.

Subscription Price

	U. S. AND CANADA	ELSEWHERE
To members of scientific societies	$5.00	$5.70
To all others	7.00	7.70

Back Number Prices

Complete set: Vol. 1, July, 1931–Vol. 16, 1945—$101.50
Yearly back number rate: $7.70
Single copies: $0.70 each.

Changes of address and proofs should be addressed to the Publications Manager.

Advertising rates supplied on request. Orders, advertising copy and cuts should be sent to the American Institute of Physics.

Entered as second class matter January 22, 1937, at the Post Office at Lancaster, Pennsylvania, under the Act of March 3, 1879. Accepted for mailing at the special rate of postage provided for in the Act of February 28, 1925, authorized May 2, 1932.

Copyright 1946, by the American Institute of Physics.

Journal
of
Applied Physics

Volume 17, Number 6 June, 1946

A Study of Distortion in Electron Microscope Projection Lenses*

JAMES HILLIER
RCA Laboratories, Princeton, New Jersey
(Received January 18, 1946)

The origin of distortion in electron microscope projection lenses is discussed and the serious nature of its effect on the measurement of particle size distributions is pointed out. Methods of measuring distortion are described. By means of first-order theory it is shown to be possible to correct distortion by the use of a two element projection lens. The degree of correction obtainable is shown to be satisfactory for most practical purposes. A double gap projection lens polepiece and the correction of distortion obtained with it are described.

I. INTRODUCTION

IN light optics it is possible to minimize the aberrations of a particular system by the appropriate combination of simple lenses. In precision electron optics, on the other hand, this has not yet been accomplished. Fortunately, while the aberrations of a simple electron lens are of the same order as those of the corresponding light optical lens, it is possible to take advantage of the very short wave-length associated with the electron beam and minimize the aberrations by radically reducing the aperture of the lens. Thus, even with a simple electron lens it is possible to produce high quality images with resolving powers of approximately 1×10^{-7} cm.

An electron microscope capable of this resolving power more than satisfies most of the immediate needs of workers interested in the fine structure of matter beyond that observable with the light microscope. This has resulted in the more concentrated efforts of electron microscope designers being directed towards the development of a practical instrument rather than to the improvement of the lens system. The fact that, in the case of the electron microscope projection lens, the aberrations produce secondary effects of a surprisingly high order of magnitude seems to have been overlooked in the literature. Some of these effects and means of correction will be discussed in the following.

Image formation by the projection lens differs in many ways from that by the objective lens. In the objective a region of the specimen, small compared with the area of the lens utilized by a single pencil,[1] is imaged by pencils of electrons having an angular aperture in the neighborhood of 3×10^{-3} radian. All points of the image are diffused by the spherical and chromatic aberrations while the off-axis points are displaced and further diffused by the additive effects of other third-order aberrations. Actually, because the

* Presented before the Electron Microscope Society Meeting in Princeton, New Jersey, November 30, 1945.

[1] The term "pencil" is used to designate those rays which take part in the imaging of a single point.

FIG. 1. Electron micrograph of a dispersion of zinc oxide pigment. In its extreme corners this image has a radial distortion (D_r) of about 30 percent. ×12,000.

diameter of the field of view is small, the last are almost negligible.[2] The image produced by the objective constitutes the "object" for the projection lens. The area of the object, in this case, is large compared with the area of the projection lens utilized by a single pencil. As the individual pencils utilized in the projection lens have an extremely small angular aperture (10^{-5} rad.) none of the third-order aberrations appears as an observable diffusion of the image. In other words, *the aberrations of the projection lens cannot reduce the resolving power of the instrument.*[3] Instead the aberrations appear as a displacement of the image points from the Gaussian or true positions; that is, as a distortion. Actually, when the projector is considered as a thin lens the conditions of imaging are such that the spherical aberration alone is responsible for this distortion. In the case of a magnetic lens there is also rotational distortion of the image but this is usually slight.

In addition to the fact that distortion in electron microscope images is undesirable from an aesthetic point of view, almost undetectable amounts can produce difficulties in the production of good stereoscopic images and can lead to completely erroneous results if the microscope is being used for the accurate determination of particle size distributions. Figure 1 is an electron micrograph of a dispersion of fine particles. This image is known to possess pincushion distortion to the extent that the radial magnification is 30 percent greater at the edge than at the center; yet it appears quite normal to the eye. Figure 2 is an electron micrograph of a regular wire screen showing the same amount of distortion. The fact that the image is distorted is apparent in this figure, though it does not appear to be excessive. The effect of this amount of distortion on a particle size distribution is illustrated in Fig. 3. Curve *A* represents the particle size distribution which would be obtained from an undistorted

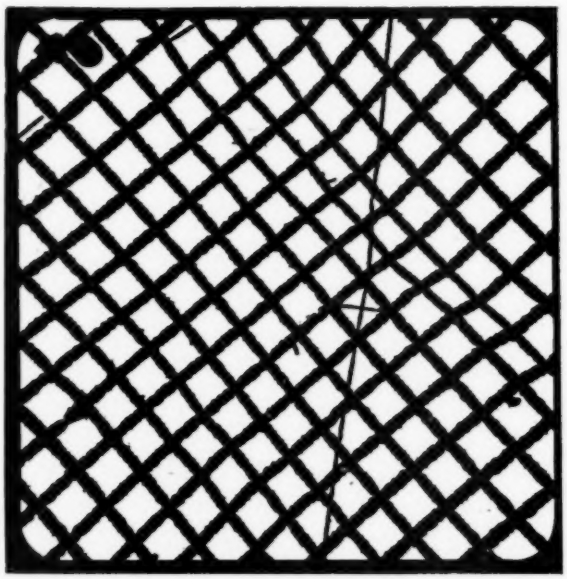


FIG. 2. A micrograph of a fairly regular wire screen made with the same projection lens and under the same conditions of distortion as the micrograph in Fig. 1.

[2] See V. K. Zworykin, G. A. Morton, E. G. Ramberg, J. Hillier, and A. W. Vance, *Electron Optics and the Electron Microscope* (John Wiley and Sons, Inc., New York, 1945).

[3] The combined effects of chromatic change in magnification and chromatic change in rotation (in the case of a magnetic lens) may be considered as an exception to the above statement. However, these affect only the outer zones of the image and are less, in general, than the defocusing of the entire image caused by the simultaneously appearing chromatic aberrations of the objective.

image of a uniform dispersion of particles which are all exactly 1000A in diameters. It is assumed that the instrument is working at a resolving power of approximately 25A. Curves *B* and *C*, on the other hand, represent the extreme distributions which would be obtained from measuring electron micrographs in which the above amount of distortion is present. Curve *B* would be obtained when particles are all measured in the circumferential direction in the image and Curve *C* when the particles are measured in the radial direction. In practice the distribution obtained would most likely lie between these two extremes, the actual curve depending on the particular arrangement used in the measurement of the images of the particles. *Thus a relatively slight amount of distortion in the electron microscope image produces a serious error in the accurate determination of a particle size distribution.*

In order to correct for such errors at present, it is necessary for the electron microscopist to calibrate the image field accurately before making critical particle size determinations. Of greater importance and convenience is the correction of the distortion by proper design of the projection lens.

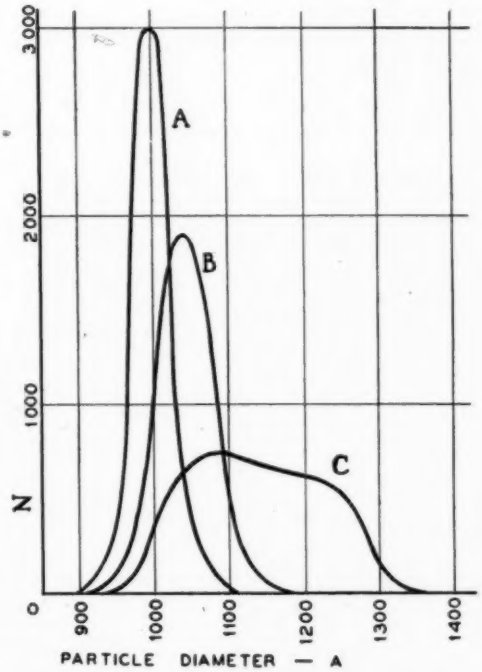


FIG. 3. A group of curves showing the effect on a particle size distribution produced by distortion of the amount present in Figs. 1 and 2.

A. Particle size distribution of a uniform suspension of particles exactly 1000A in diameter obtained with a distortion-free system.
B. Particle size distribution of the same specimen which would be obtained by measurement in the circumferential direction in an image possessing distortion ($D_r = 30$ percent at edge).
C. Particle size distribution of the same specimen which would be obtained by measurement in the radial direction.

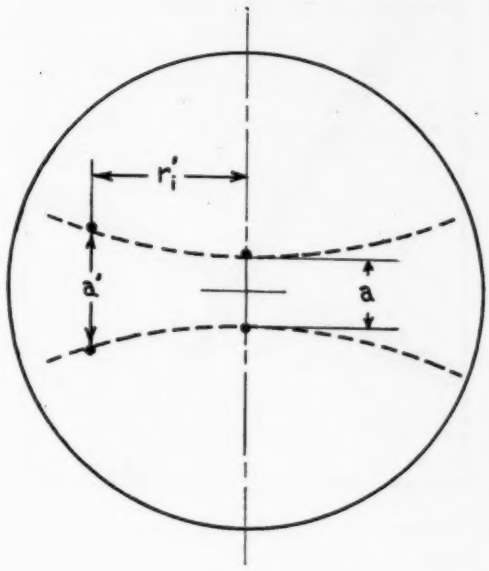


FIG. 4. Diagram showing the measurements which permit the complete determination of pincushion distortion at any point of an image.

II. MEASUREMENT OF DISTORTION

Distortion in an electron microscope image involves a displacement of the image points, usually in the radial direction, from a geometrically true representation of the specimen. Thus, if r_i', θ_i' are coordinates of the actual image point and r_i, θ_i the coordinates of the corresponding point in the geometrically true image

$$r_i' = F(r_i); \quad \theta_i' = \theta_i + G(r_i), \tag{1}$$

where $F(r_i)$ and $G(r_i)$ are continuous, single-valued functions of r_i. $G(r_i) = 0$ except in the case of the magnetic projection lens; in any case it is small and will be neglected in the present work. The expressions (1) represent the transformation of coordinates necessary to transform the true image into the distorted image. In practice, however, the distortion can be considered as a point to point variation in the magnification since that is the way in which it disturbs the accuracy of measurements made in the image. It is obvious that the variation in

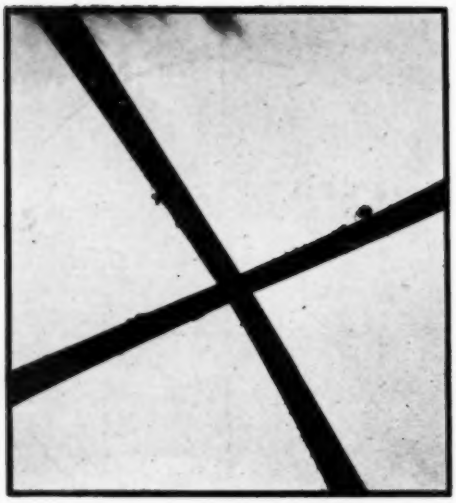


FIG. 5. An image produced by the projection lens of stretched cross wires from which it is possible to determine completely the distortion pattern of the lens.

magnification depends on the direction of measurement in the image.

Consider now two points of the image which are separated by a distance a and which straddle the central point of the image (see Fig. 4). By moving the specimen in such a way that these points move away from the center of the image in the direction perpendicular to the line joining them they will trace out two curves which, as will be seen below, give all the information necessary for the complete determination of the distortion.

It is immediately obvious that $(2\pi r_i'/a') = (2\pi r_i/a)$ and hence the circumferential distortion D_c is given by the expression:

$$(D_c/100)+1 = a'/a = r_i'/r_i = F(r_i)/r_i, \quad (2)$$

where D_c is defined as the percentage difference between the magnification at the point under consideration (measured in the circumferential direction) and that at the center of the image. For high accuracy a must be small.

The radial distortion D_r which is similarly defined with the exception that the measurements are made in the radial direction is given by the expression:

$$\frac{D_r}{100}+1 = \lim_{r_{i2}-r_{i1}\to 0}\left(\frac{r_{i2}'-r_{i1}'}{r_{i2}-r_{i1}}\right) = F'(r_i). \quad (3)$$

The actual procedure of measuring the distortion is now quite straightforward. The necessary data are obtained by making a series of micrographs in which a selected region of the specimen is moved in the way indicated in Fig. 4. If the projection lens of the instrument is easily accessible, it is more convenient to arrange two cross-wires in a holder just above it and make a single micrograph as shown in Fig. 5. Ordinary finely drawn wires can be depended on to have the same diameter over a considerable length. In either case it is possible to obtain precise measurements of (a'/a) and r_i' from which a curve can be drawn showing the relationship between D_c and r_i'. From these measurements it is also possible to calculate the corresponding values of r_i and plot r_i' as a function of r_i. The value of $[(D_r/100)+1]$ can then be determined for any point of the image by taking the slope of the curve for r_i' as a function of r_i at the desired points and plotting against r_i'. Figure 6 gives two curves for a typical projection lens system working at a low magnification for which the excessive distortion at the edge of the image is visible. From such measurements a calibration chart can be drawn up which gives the radial and

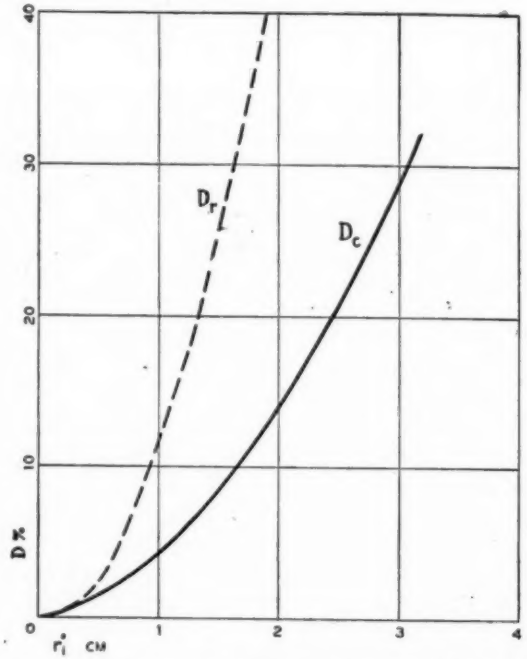


FIG. 6. Radial (D_r) and circumferential (D_c) distortion curves for an early projection lens polepiece working at a low magnification for which the excessive distortion at the edges of the image is visible.

circumferential distortion for each zone of the electron microscope image. If, in particle size measurements, a standard electronic magnification, optical magnification, and measuring technique is used, it is a simple matter to utilize such a chart to compensate for the effects of image distortion. It is important, of course, that the electron-optical center of the projection lens image be accurately known, if true correction is to be obtained. This is normally determined at the same time as the calibrating measurements are made.

III. CORRECTION OF DISTORTION IN A SINGLE LENS

Incidental to the study of distortion of magnetic projection lenses, it was observed that the distortion is dependent on the spacing of the polepieces, being reduced as the spacing is increased. The effect is also predicted from a theoretical study of spherical aberration.[2] The experimental curve shown in Fig. 7 illustrates this effect. While the improvement attainable is appreciable it is accompanied by considerable loss of magnification. The method is of some practical and academic interest, but it does not represent a complete solution of the problem.

IV. TWO-LENS CORRECTION OF DISTORTION

It was realized a few years ago that the use of a double projection lens with the first lens situated *between the second lens and its object side focal point* provided some correction of image distortion if the strength of the two lenses were carefully adjusted. The qualitative idea was tested and the results described at the time,[4] but it was not investigated further because of the apparent necessity of using a specially constructed double projection lens coil and an extra regulated current supply. Figure 8 shows the initial results obtained while Fig. 9 is a ray diagram explaining the correcting action of the second lens.

In the present work a mathematical study of such a system has been carried out to show that it is possible to obtain any desired degree of correction with a properly adjusted two lens system and that in practice it is possible to design a distortion-free polepiece system actuated by a single lens coil.

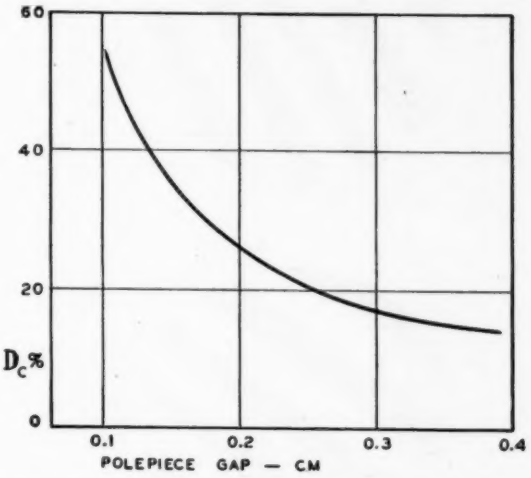


FIG. 7. Curve showing the effect on circumferential distortion of increasing polepiece gap. The curve was taken at constant excitation of the lens, and hence there is an accompanying change in magnification. By drawing the curve for a fixed position in the intermediate image, the effect of this change in magnification is eliminated.

V. THEORY OF TWO-LENS CORRECTION

Figure 10 is a ray diagram indicating the significance of the symbols used in the following. The system consists of two thin lenses of paraxial focal lengths f_{10} and f_{20} separated by a distance d_1. The image produced by the objective and remagnified by these lenses consists of pencils essentially parallel to the axis and of extremely small angular aperture (10^{-5} rad.). For all practical purposes these pencils can be considered as single rays leaving a distant point.[5]

Since pincushion distortion only is being considered the problem involves determining the characteristics of the second lens which will make r_i' a linear function of r_1. From the geometry of Fig. 10 it can be seen that

$$r_2 = r_1(f_1 - d_1)/f_1, \tag{4}$$

$$r_1' = r_2(d - d_1 - v_2)/v_2. \tag{5}$$

Then if the simple lens formula $(1/v)-(1/u)=(1/f)$ is assumed to hold for both lenses (5)

[4] J. Hillier and R. F. Baker, "Miscellaneous research on electron microscope parameters," presented before Electron Microscope Society of America, January 14, 1944.

[5] For simplicity this point will be considered at infinity in the present paper. In critical work, however, it is necessary to take into account the fact that the objective is normally at a finite distance from the projector in which case u_1 is substituted for f_1 in the following formulas.

FIG. 8. Projection lens images of parallel stretched wires showing correction of distortion obtained in early work on two-lens projection systems. (a) Uncorrected. (b) Corrected.

becomes, after multiplying and collecting terms:

$$r_i' = r_1\left[(d-d_1)\left(\frac{1}{f_1}+\frac{1}{f_2}\right)-\frac{d_1(d-d_1)}{f_1 f_2}+\frac{d_1}{f_1}-1\right]. \quad (6)$$

Because of spherical aberration f_1 and f_2 depend on r_1 and r_2, respectively, according to the relations

$$f_1 = f_{10}(1-k_1 r_1^2) \quad \text{and} \quad f_2 = f_{20}(1-k_2 r_2^2), \quad (7)$$

where k_1 and k_2 are the spherical aberration constants.[6]

Substituting the expressions (7) for f_1 and f_2 in (6) leads to

$$r_i' = r_1\left[(d-d_1)\left\{\frac{1+k_1 r_1^2}{f_{10}}+\frac{1+k_2 r_2^2}{f_{20}}\right\}-\frac{d_1(d-d_1)}{f_{10}f_{20}}(1+k_1 r_1^2+k_2 r_2^2)+\frac{d_1}{f_{10}}(1+k_1 r_1^2)-1\right], \quad (8)$$

where $k_1 r_1^2$ and $k_2 r_2^2$ are considered small compared to 1 in view of the experimental values obtained for these quantities. Collecting coefficients of $k_1 r_1^2$ and $k_2 r_2^2$ results in an expression of the form

$$r_i' = r_1(A_0 + A_1 k_1 r_1^2 + A_2 k_2 r_2^2). \quad (9)$$

Hence to make r_i' a linear function of r_1, it is necessary to equate the coefficients of the higher

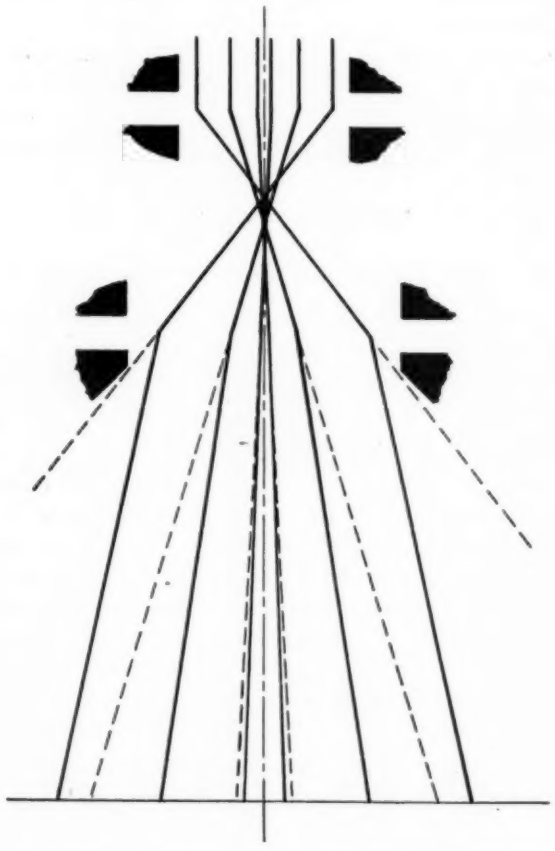


FIG. 9. Ray diagram showing how the combined effects of focusing and spherical aberration in a weak additional component of projection lens system can correct for pincushion distortion. The focal point of the lower lens is situated some distance above the plane of the upper lens.

[6] The constant k defined above is more convenient for discussing spherical aberration in electron lenses than the commonly used spherical aberration coefficient C. C is a dimensionless constant for a lens of *fixed focal length*. However, one of the niceties of electron optics is the fact that the focal length of a lens is variable over wide ranges, in which case the use of C becomes cumbersome and leads to difficulties in the interpretation of experimental results. The constant k, on the other hand, is independent of the focal length provided that the lens in question may be regarded as a thin lens. The two coefficients are related by the equation $kf^3 = Cf$.

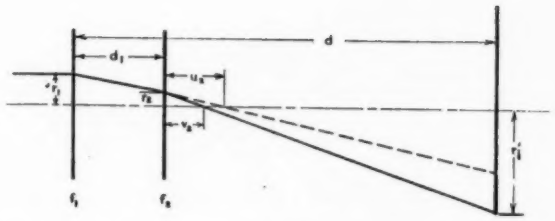


FIG. 10. Ray diagram showing the significance of the symbols used in the text.

powers of r_1 and r_2 to zero, that is

$$A_1=\frac{d-d_1}{f_{10}}-\frac{d_1(d-d_1)}{f_{10}f_{20}}+\frac{d_1}{f_{10}}=0, \tag{10}$$

$$A_2=\frac{d-d_1}{f_{20}}-\frac{d_1(d-d_1)}{f_{10}f_{20}}=0. \tag{11}$$

Solving (10) and (11) for f_{10} and f_{20} gives

$$f_{10}=d_1, \tag{12}$$

$$f_{20}=\frac{d_1(d-d_1)}{d}=\frac{f_{10}(d-f_{10})}{d}. \tag{13}$$

Thus the above has shown that a first-order correction of distortion can be made by the use of two lenses where the second lens is placed at the focal point of the first lens and has a focal length given by (13). Since $f_{10}\ll d$ the focal lengths of the two lenses are very nearly equal. It should also be noted that the conditions for correction are independent of the spherical aberration of either lens.

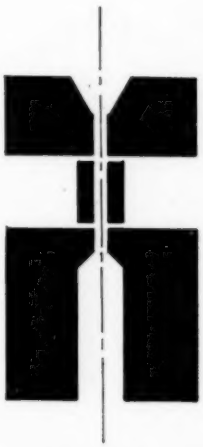


FIG. 11. The schematic cross section of a double gap projection lens polepiece which was used in the present experiments.

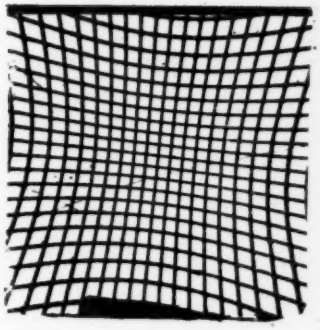


FIG. 12a. The distortion pattern obtained with a single section of the double polepiece section activated.

FIG. 12b. Distortion pattern obtained with both sections of the double lens polepiece activated.

To investigate the degree of correction attained by the above method it is necessary to return to the complete expression (5) which on substituting for u_2 and making use of (4) becomes

$$\frac{r_i'}{r_1}+1=\frac{d-d_1}{f_2}\cdot\frac{r_2}{r_1}+\frac{d}{f_1}. \tag{14}$$

In order to simplify the notation

$$(r_2/r_1)=(f_1-d_1)/f_1$$

will be written as ϵ. Then (14) may be written as

$$\left(\frac{D_c}{100}+1\right)M+1=\frac{(d-d_1)\epsilon}{f_2}+\frac{d(1-\epsilon)}{d_1}, \tag{15}$$

where $M=(r_i/r_1)$ is the undistorted magnification of the system. If the values for the corrected system are introduced it is also possible to write

$$r_1^2=-\epsilon/[k_1(1-\epsilon)] \tag{16}$$

and

$$r_2^2=-\epsilon^3/[k_1(1-\epsilon)], \tag{17}$$

whereupon the complete expression for f_2 becomes

$$f_2=\frac{d_1(d-d_1)}{d}\left[1+\frac{k_2\epsilon^3}{k_1(1-\epsilon)}\right]. \tag{18}$$

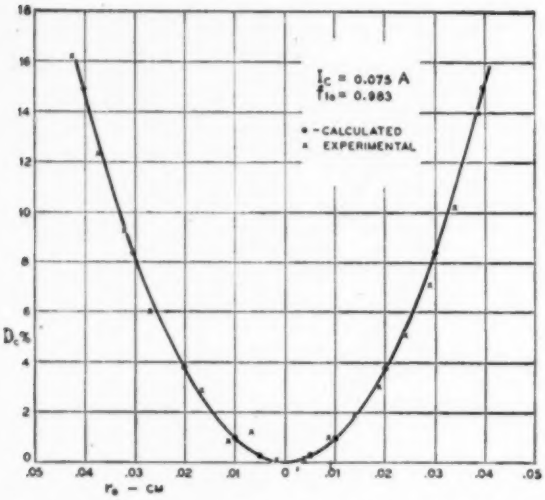


FIG. 13. A comparison between the calculated and experimental distortion pattern for a single section of the polepiece shown in Fig. 11.

Substituting (18) in (15) and rearranging gives

$$\frac{D_e}{100}=-\left(1+\frac{1}{M}\right)\times\left[\frac{k_2}{k_1}\cdot\frac{\epsilon^4}{1-\epsilon}\Big/\left(1+\frac{k_2}{k_1}\frac{\epsilon^3}{1-\epsilon}\right)\right]. \quad (19)$$

Since $\epsilon=-k_1r_1^2/(1-k_1r_1^2)$ and $k_1r_1^2\ll 1$ the expression (19) shows that a system corrected by the application of the relations (12) and (13) provides an image which deviates from the true image only in the fourth power of the small quantity $k_1r_1^2$.

Before leaving the theoretical discussion, it should be pointed out that even more complete correction can be obtained under the present assumptions if only the condition (12) is imposed on (15) and the expression

$$\left(\frac{D_e}{100}+1\right)M+1=\frac{d}{d_1}+\epsilon\left\{\frac{d-d_1}{f_{20}\left[1+\frac{k_2}{k_1}\frac{\epsilon^3}{1-\epsilon}\right]}-\frac{d}{d_1}\right\}. \quad (20)$$

In this case a more favorable value of f_{20} can be obtained by making the coefficient of ϵ equal to zero for ϵ equal to its maximum possible value ϵ_{max} in a given practical example, that is by taking

$$f_{20}=\frac{d_1(d-d_1)}{d\left[1+\frac{k_2}{k_1}\frac{\epsilon^3{}_{max}}{1-\epsilon_{max}}\right]}. \quad (21)$$

VI. EXPERIMENTAL RESULTS

The realization in practice of a projection lens system which fulfills the above derived conditions is not difficult if it is possible to excite the individual lenses independently. Unfortunately, this requires a rather elaborate special coil structure and an additional regulated source of current if it is to be easily adaptable to existing electron microscopes. On the other hand, if the two lenses are made part of one magnetic circuit excited by a single coil, the proper ratio between the focal lengths of the lenses must be attained by the exact adjustment of the physical shapes and dimensions of the lenses. Owing to the lack of accurate methods of computing the relative field strengths in the lenses of such an arrangement that adjustment can be accomplished only by tedious empirical methods. For these reasons a direct experimental test of the above theory was not attempted in the present work. Instead, a simple special case which could easily be achieved in practice and to which the essential formulas could be applied was selected.

An experimental polepiece assembly was constructed and tested in a demountable electron-optical system. This assembly, which is shown schematically in Fig. 11, included two mechanically identical magnetic lenses separated by 0.9 cm, giving maximum correction of distortion under conditions corresponding to a magnification of approximately 4000× in an RCA Type B electron microscope. The flat-face polepiece shape for the two lenses was chosen because it was known from previous work that such a shape gives particularly high distortion.

As a first test, the lower lens was inactivated and the distortion pattern of the upper lens obtained for a series of different excitations. Figure 12a is one of the patterns obtained. An accurately made electrolytic copper screen (500 mesh) was used to determine the distortion patterns of the system. In Fig. 13 the circumferential distortion D_e is plotted against the radial position in the object (r_0). In the

experimental arrangement both the source and the test specimen were at finite distances from the lens. Taking account of this, a calculation of the value of the spherical aberration constant k_1 for a point at the extreme edge of the image yields a theoretical curve for D_c as a function of r_0. It can be seen that the agreement between the experimental and computed curves is exact to within the experimental error. This justifies the use, in practice, of the expression $f_1=f_{10}(1-k_1r_1^2)$ for the focal length of the lens. The value of k_1 obtained was 42.5 cm^{-2} which, as expected, was much higher than that for a normal projection lens (6–8 cm^{-2}). Even with this high value for k_1, ϵ^4 in (19) has a maximum value of only 10^{-4}. The value of k_1 was found to be independent of lens strength except at high values of the lens current. This gives further experimental justification of the use of this coefficient in electron optics.

Using the system with both lenses activated the distortion patterns for the same series of lens excitations were obtained. The pattern showing maximum correction of distortion is given in Fig. 12b. In the right-hand side of Fig. 14 experimental values of D_c are plotted against r_0 for four values of excitation near the point of maximum correction. Four calculated curves using $k_1=k_2=42.5$ cm^{-2} and the indicated values of $f_{10}=f_{20}$ are plotted in the left-half of Fig. 14. The correspondence is seen to be quite good. From the geometrical arrangement of the apparatus and from the magnification at the center of the image it is possible to calculate the values of the focal lengths of the lenses if they are assumed to be equal. However, on carrying out the calculations it was found that the values of f_{10} obtained did not agree with those expected from the calculated curves. The discrepancy which was greater than the experimental error was probably partly caused by the fact that the leakage fields in the coil system used destroyed the equality of the focal lengths of the two lenses and partly by the fact that the lenses used were not strictly thin lenses.

It should be pointed out in passing that some rotational distortion is present in the pattern of Fig. 12, but that it is of negligible magnitude.

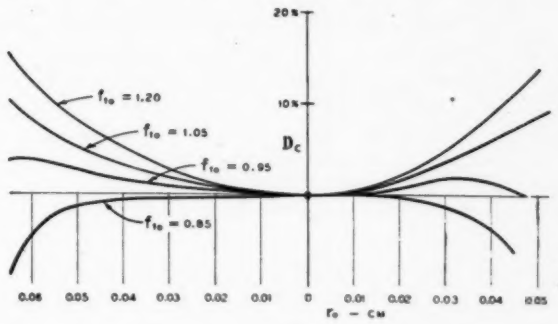


FIG. 14. Comparison between calculated and experimental values of the circumferencial distortion for the double lens polepiece. In the left half, calculated values for the designated focal lengths of the lenses are given. In the right-hand half a corresponding series of experimental curves is given.

CONCLUSIONS

It has been shown in the above discussion that relatively small amounts of distortion in the projection lenses of electron microscopes can produce serious disturbances in particle size distributions obtained with the instrument. Furthermore, it has been shown that it is both theoretically and practically possible to design a projection lens system for an electron microscope in which, at least at one magnification, the distortion due to spherical aberration of that lens is reduced to negligible proportions.

ACKNOWLEDGMENTS

The author gratefully acknowledges his indebtedness to R. F. Baker and J. M. Morgan who made many of the earlier and more tedious measurements of distortion. He also wishes to thank G. W. Brown for many helpful discussions.

Parafocusing Properties of Microcrystalline Powder Layers in X-Ray Diffraction Applied to the Design of X-Ray Goniometers

J. C. M. Brentano
Northwestern University, Evanston, Illinois

(Received December 17, 1945: Revised February 25, 1946)

A distinction is drawn between the focusing properties of single crystals with regard to x-ray diffraction and the parafocusing properties of microcrystalline powder layers. A discussion of these parafocusing properties and of the absorption term, which arises in the diffraction of x-rays from powder layers, is followed by a discussion of various types of powder goniometers. Their characteristic properties are compared. The application of parafocusing conditions to the design of a goniometer for quantitative evaluations and of back reflection cameras is discussed in greater detail.

SOME recent developments of structure analysis and the extension of the identification method in the direction of quantitative evaluations have stressed the demand for the quantitative measurement of the relative intensities of x-ray reflections from microcrystalline powders and have given greater significance to the so-called focusing methods in which the effect of absorption on relative intensities can be eliminated. The larger angular apertures used in these arrangements as compared with the original powder rod camera is also of particular interest in industrial applications where the reduction of exposure times is relevant. While the Debye-Scherrer-Hull method, which has the inherent merits of simplicity and of utilizing minute quantities of material has been largely standardized, the requirements mentioned above have given rise to a variety of methods. It is proposed to discuss in the present paper[1] some basic points which arise with these methods with the object of bringing out their characteristics and possibly furthering the development of arrangements best suited to meet specific conditions.

1. X-RAY DIFFRACTION FROM POWDERS

The term focusing is sometimes applied to two distinct conditions of x-ray diffraction. Without wishing to break with the widely accepted use, it may be helpful for the present discussion to set out the distinction between real focusing and the ray-collective properties of powder arrangements which more appropriately can be described as parafocusing. The distinction has not only bearing on the underlying geometrical-optical relations, with which we are here mainly concerned, but also on the purpose for which one and the other type of focusing can effectively be used.

The term focusing as defined in geometrical optics is related to the existence of an object and image space in such correlation that to a ray in the object space corresponds a ray in the image space and that to a point in the object space, defined by the intersection of a bundle of rays, is correlated a point in the image space in which the corresponding rays intersect. In x-ray optics of single crystals such a correlation can be established for monochromatic beams of wavelength λ on the ground of a "reflection" from a set of lattice planes of spacing d satisfying the Bragg law:

$$\sin\theta = n\lambda/2d. \tag{1}$$

Examples for this are the single bent crystal, the Cauchois monochromator, and the Bragg single crystal goniometer, all having focusing properties though limited to beams in one plane. No such image to object relation applies to the ray-collective arrangements of crystal powders, which may be described as parafocusing. Parafocusing is concerned with the geometrical locus of the powder particles and by disregarding their orientation establishes a necessary, not a sufficient, condition. A closer reference to the Bragg single crystal goniometer and to the parafocusing property of a flat powder layer may be helpful in bringing out the difference. In the Bragg single crystal goniometer the face of an

[1] Paper given at the Ann Arbor Summer Meeting of the American Society for X-Ray and Electron Diffraction, 1943. Abstract in Phys. Rev. **64**, 312 (1943).

extended crystal is so placed as to comprise the axis of rotation C (Fig. 1) of the crystal, when two points A and B in a plane normal to the axis at equal distance from C with the angle $ACB=\pi-2\theta$ are so correlated that when the crystal is rotated various points of its surface are brought in succession in such positions as to reflect the radiation incident in the plane ACB according to (1) from A to B. The whole crystal surface does not reflect simultaneously, but true focusing is satisfied in the sense that whenever the particular reflection defined by (1) occurs, the reflected ray will go through the same point B. The flat crystal face which comprises the axis C satisfies this condition exactly. Comparing this with the conditions which arise if we substitute for the single crystal a flat layer of randomly oriented microcrystalline particles, we find that the relation peculiar to the Bragg single crystal goniometer, which limits the reflection to the particular crystal setting for which it will be directed to B, does not apply, but that for each volume element of the powder the reflection (1) will take place according to the surface of a cone of semi-aperture 2θ, having the incident beam as axis. Of this diffracted radiation some will be received at B in so far as a generatrix of the diffraction cone exists which goes through B. The fraction of the radiation so diffracted will always be very small. Various volume elements will give additive contributions to the reflection (1) received at B if generatrices of the diffraction cones for the various elements intersect in B; for the points of a flat surface this is satisfied only in approximation, the correct surface for the powder being a surface of double curvature. This additive effect then takes place simultaneously for all parts of the powder and the best superposition is obtained for one particular setting of the layer. While the relation for the true focusing from a single crystal can only be satisfied when the distances of A and B from C are equal, the approximate parafocusing condition established by the writer for a flat powder layer, relation (3) of Section 6, applies equally well for any ratio of distances AC and BC. The intensity relations are also entirely different. For true focusing from the single crystal the intensity of the reflected radiation is independent of the crystal size and angular aperture of the incident beam so long as the angular range, through which the crystal is rocked, does not exceed the range within which some part of it is reflecting; for the crystal powder for which all parts of the layer contribute simultaneously to the reflection, the intensity increases with the aperture of the incident beam and the area of the crystal surface. The geometry of the reflection from the single crystal rotated through an extended angular range provides that, when true focusing is satisfied, the reflections for different angles 2θ resulting for different λ's or different d's are recorded on a photographic film on a cylindrical surface along AB as sharp marks; while with the crystal powder, in recording an extended angular range, blurring occurs whether the powder is kept stationary or whether it is rotated unless a special screening diaphragm is used.[2]

In the true focusing from a single crystal, a large fraction of the reflected radiation can be directed into one particular direction or into one particular point; it is used in single crystal spectrometers and monochromators. In the parafocusing of crystal powders the reflection from the individual powder element is used at very low efficiency. It satisfies the purpose of increasing the intensities of reflections by adding the contributions from a large volume; it further makes it possible to establish powder layers which satisfy very simple absorption conditions and to average over large powder volumes. This is of significance in the quantitative evaluation of intensities and in the determination of the contour of diffraction lines.

2. PARAFOCUSING TOROIDAL SURFACE

The surface, which represents the geometrical locus of all points in space so situated that for a point source in A (Fig. 2) a ray deflected through a given angle 2θ can go through a point B, is obtained by the rotation of an arc with the

[2] Sir William Bragg (Proc. Phys. Soc. **33**, 222 (1921)) was able to observe the reflections from a powder layer with an ionization goniometer without taking account of the difference between the two kinds of focusing when the entrance slit to the ionization chamber acted as a screening diaphragm. Experiments of photographic recording, done in his laboratory with fatty acids, were successful (A. Muller, J. Chem. Soc. **123**, 2043 (1923)) because these presented preferential orientation and were equivalent to single crystals, but failed to give well-defined lines for a randomly oriented powder.

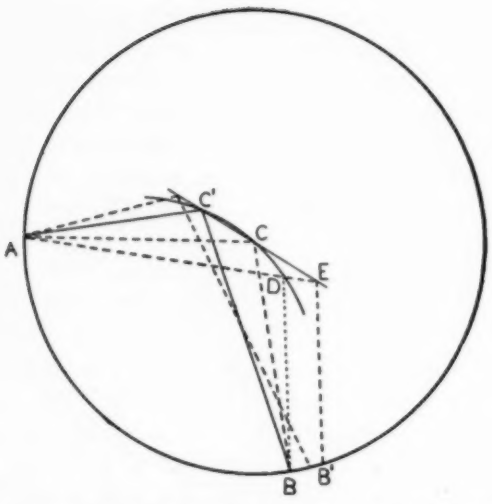


FIG. 1. The straight line through point C and C' indicates a particular setting of the face of a single crystal or of the surface of a flat powder layer going through the axis of the camera. The full line $AC'B$ indicates the reflection from the single crystal for that particular setting: as the crystal is rotated reflection takes place from various points in succession and according to Bragg's focusing relation is always directed to the same point in B. With the powder the same reflection for the fixed position occurs simultaneously from various points of the surface as indicated by the dotted lines without the Bragg focusing condition being satisfied. The curved line through $C'C$ and D marks the intersection of the plane ACB with the parafocusing toroidal surface. The projection BB' of the distance DE measures the aberration for the particular ray as discussed in Section 2.

peripheral angle $ACB=\pi-2\theta$ about the chord AB. This surface does not satisfy the condition of focusing, but only of parafocusing: it does not establish the orientation but only the location of the crystal elements which, for a given spacing and wave-length, may contribute to the deflection of radiation from the source A to B on the ground of relation (1). This surface has been discussed by the writer in an earlier paper,[3] where it was pointed out that the orientation of the lattice planes, which actually contribute to the diffracted radiation in B, is such that their normal bisects angle ACB so that they are inclined to the toroidal surface, except for the median annular zone in C'. This is the orientation given to the reflecting lattice planes in the bent crystal monochromator. The angle 4θ between the tangents to the surface in A determines the largest possible aperture of a beam from A which can be collected in B by one deflection through the angle 2θ. With single crystals only equatorial rings or longitudinal bands tangential to the parafocusing surface can be made while with a powder it is not practical to utilize the full extent of this surface in an instrument intended to cover a range of deflection angles. Reference to this parafocusing surface, however, is useful in evaluating the efficiency of powder arrangements which approximate parts of the parafocusing toroid and offers a convenient way for determining the aberrations which result with the actual surface used. The aberration for any particular point of the powder layer is obtained by projecting its distance from the corresponding point of the toroidal surface on the plane in which the diffraction pattern is recorded. This is indicated by Fig. 1 in which DE is the distance between corresponding points of the toroid and powder element for the incident ray ADE and BB' measures the aberration. This way of evaluating aberrations saves considerable time compared with direct calculation.

The contribution in B resulting from any particular part of the toroid when represented by the surface of a powder layer is further determined by an absorption term in the intensity expression. For a narrow beam directed to a surface element of a powder layer with the absorption coefficient μ, which is sufficiently thick to be all absorbing, the absorption term is[4]

$$\frac{1}{\mu(1+\sin\alpha/\sin\beta)}, \tag{2}$$

where α and β are the glancing angles of incidence and of reflection.

[3] J. C. M. Brentano, Arch. Sci. Phys. et Nat. [4] **44**, 66 (1917).

[4] J. C. M. Brentano, Phil. Mag. **4**, 620 (1927); **6**, 178 (1928); Proc. Phys. Soc. **47**, 932 (1935). For an all-absorbing layer the full expression for the diffracted intensity P_b received at the distance b by the entrance slit of an ionization chamber of unit height and such width as to comprise the whole reflection is

$$P_b=\frac{I p_{(hkl)}F^2_{(hkl)}N^2 q_{(hkl)}f}{8\pi b\sin\theta\,\mu(1+\sin\alpha/\sin\beta)},$$

while the corresponding expression for a small volume element dV with negligible absorption is

$$dP_b=\frac{I p_{(hkl)}F^2_{(hkl)}N^2 q_{(hkl)}dV}{8\pi b\sin\theta}.$$

In these expressions I is the intensity of the incident beam per unit cross section and f is its cross section at the powder. The other symbols have the usual meaning as used in the last paper quoted. In this paper a misprint has actually occurred in the expression equivalent to P_b ((5) in the paper); the expressions related to it are given correctly in the paper.

3. REQUIREMENTS WHICH ARISE IN THE EVALUATION OF X-RAY POWDER DIFFRACTION PATTERNS

For *crystal structure analysis* single crystals present the advantage of giving directly the correlation between the individual reflections and the orientation of the reflecting lattice planes in the crystal. Powder diffraction methods are resorted to, apart from such instances where large single crystals are not available, when it is desired to obtain relative intensity data free from extinction from powder particles for which the size and regularity of the microcrystalline domains is so reduced as to make primary extinction small. Powder methods in the form of the quantitative powder admixture[4] provide a simple way for interrelating the intensities of two constituents of a mixed powder to the scattering factors and to the ratio of the masses of the constituents and for obtaining absolute intensities of reflections by the admixture of a standardizing substance, for which the unit cell scattering factor is known. The intensities of the reflections to be studied can then be related to those of the standardizing admixture; they can thus be expressed in absolute measure without the necessity of measuring the intensity of the incident beam and to correct for absorption.

Crystal structure analysis requires (a) that one single record should cover an extended angular range, (b) that the deflection angles should be recorded to a high degree of accuracy and preferably that line distances should be proportionate to angular distances so as to facilitate interpolations, and (c) that relative integrated intensities of reflections for different angles should not be affected by the shape and absorption of the particular specimen. It is further sometimes required (d) to obtain diffraction patterns with very small quantities of material, while in another group of measurements it is desired (e) to make use of large powder volumes. This last requirement arises when the powder contains large crystalline domains of say more than 10^{-5} cm diameter when it is desired to suppress the "spottiness" of lines which arises under such conditions when the diffraction pattern is obtained from a small powder volume.

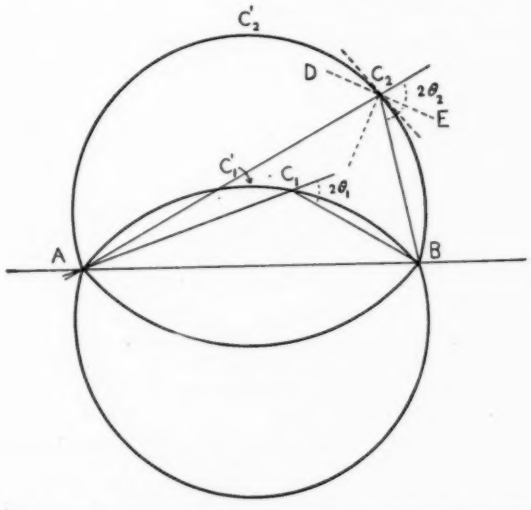


FIG. 2. Lines $AC_1'C_1B$ and $AC_2'C_2B$ indicate the intersections of two parafocusing toroidal surfaces for the angles $2\theta_1$ and $2\theta_2$ with the plane of drawing. The dotted line DE indicates the orientation of the diffracting lattice planes for the point C_2, the normal of which bisects angle AC_2B.

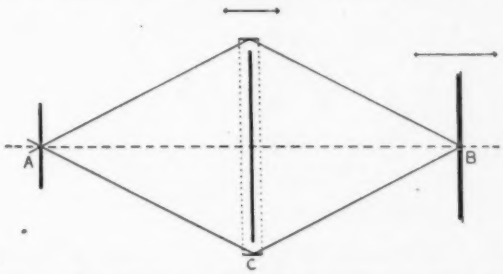


FIG. 3. Camera using an annular powder layer. For a fixed position of the ring each point of the axis is associated with a definite deflection angle.

For *evaluating the degree of regularity and the size distribution of microcrystalline domains* from the line width, or better from the line contour, it is essential that (f) the line contour should not be affected by the shape and absorption characteristics of the specimen or that it should be possible to allow for such effects. It is then of particular significance to suppress irregularities of intensity distribution produced by the presence of large crystallites as indicated under (e). This implies the use of large powder volumes, incident beams of large aperture and rotation of the specimen.

For *the method of identification*, in which chemical constituents or phases are identified by their characteristic x-ray diffraction patterns in the form developed by Hanawalt, Rinne, and Frevel and by Davey,[5] the requirements are

[5] J. D. Hanawalt, H. W. Rinne, and L. K. Frevel, Ind. Eng. Chem. Anal. Ed. **10**, 457 (1938) and W. P. Davey, J. App. Phys. **10**, 820 (1939).

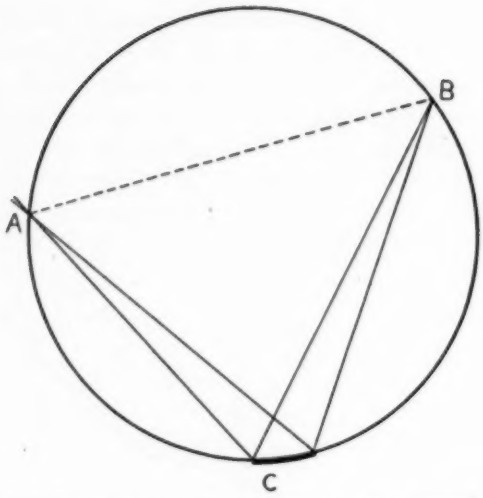


FIG. 4. Bohlin-Seemann type camera. The powder in C forms part of a cylindrical surface which approximates toroids constructed by rotating arc ACB about cord AB where A is the source and B for different θ's is located at various points of the circle.

similar to those of structure analysis, only to a less exacting degree.

In *technical applications*, whether they comprise the recording of line contours or of identification patterns, the speed of recording, which is only a minor factor in structure analysis, assumes major significance. With the Debye-Hull camera and photographic recording the exposure times are longer than the times required for spectroscopic analyses or for simple chemical tests. In the field of identifications x-ray methods are thus limited to instances where the more rapid methods are inadequate, e.g., the distinction between different morphological phases or between constituents which cannot easily be separated when present in a mixture.

By using counters or ionization chambers instead of photographic recording the sensitivity can be considerably increased. It should be noticed, however, that applied to the recording of weak reflections the process of point to point scanning used with these methods necessitates slowing the scanning rate sufficiently to secure that the count numbers with which the intensities for the individual angular elements are recorded are great enough to suppress large statistical fluctuations. This limits the speed at which such measurements can be taken, particularly when the intensities are small. Industrial applications using photographic, counter, or ionization recording set thus the demand for utilizing beams of high intensity and large angular aperture so as to reduce the exposure times.

4. THE DEBYE-SCHERRER-HULL CAMERA

The demands set out in the previous section may be considered in relation to the classical Debye-Scherrer-Hull camera. This gives patterns covering an extended angular range with proportionality between angles and line distances. The line width so far as resulting from collimating conditions is of the order of the diameter of the x-ray beam incident on the powder specimen. Very sharp lines can thus be obtained at the expense of exposure time by the use of collimating pinholes of small diameter. The effect of absorption introduces some complication, both in the evaluation of spacings and of intensities. A discussion of these conditions has been given in a paper by Möller and Reis[6] and specific methods have been developed by McKeehan,[7] Claassen,[8] Bradley and Jay,[9] and others. Recently a detailed analysis of such corrections has been given by Warren.[10] Powder rods which simplify the allowance for absorption have been proposed by Greenwood[11] and by Rusterholz.[12] The limi-

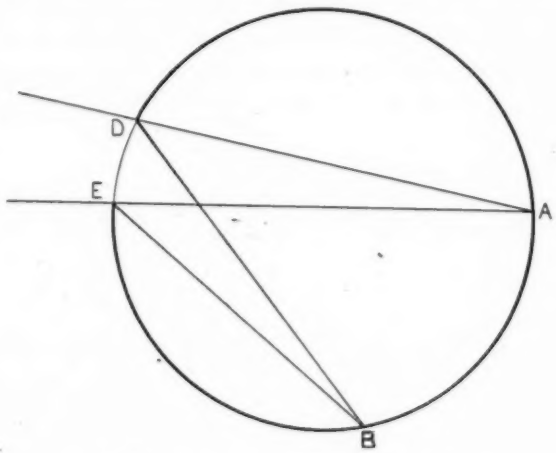


FIG. 5. Guinier type camera. The incident beam passes through the powder layer DE and converges toward point A. The diffracted beams are collected at points B of a cylindrical surface.

[6] H. Möller and A. Reis, Zeits. f. physik. Chemie **A139**, 425 (1928).
[7] L. W. McKeehan, J. Frank. Inst. **193**, 231 (1922).
[8] A. Claassen, Phil. Mag. [7] **9**, 57 (1930).
[9] A. J. Bradley and A. H. Jay, Proc. Phys. Soc. **44**, 563 (1932). A. J. Bradley, Proc. Phys. Soc. **47**, 879 (1935).
[10] B. E. Warren, J. App. Phys. **16**, 614 (1945).
[11] G. Greenwood, Phil. Mag. [7] **3**, 963 (1927).
[12] A. Rusterholz, Helv. Phys. Acta **4**, 68 (1931).

tation of these procedures is that they depend on making use of a powder rod of definite shape and definite absorption, which is difficult to obtain. The effect of absorption in modifying the relative intensities can be made small by using a powder rod or wedge of negligible absorption, either by reducing its thickness or by diluting it and by the use of penetrating radiation. It will be seen that under such conditions, which reduce absorption effects, the exposure times become excessively long. In practice a compromise is made between the elimination of absorption errors and the permissible exposure times. With very thin or diluted specimens the effect of large crystallites in producing irregular lines is accentuated. It can be reduced by displacing the powder rod longitudinally or rotating it transversally.[13]

The Debye-Hull camera requires a minute quantity of material and permits to correct for faulty centering of the specimen.[14] Detailed descriptions of cameras have been given by Bradley, Lipson, and Patch[15] and recently by Buerger.[16] Several commercial types are also available.

It will be seen that this camera can satisfy most of the demands of structure analysis and of line contour measurements and this should be considered in comparing it with other arrangements described further on. Its limitations are the long exposure times and the difficulty of making correct allowance for absorption in determining relative intensities taken over a wide angular range. When using small or diluted specimens to reduce the absorption effect the line contours are sensitive to the presence of large crystallites.

5. CAMERAS USING CYLINDRICAL POWDER STRIPS

Approximations to the toroidal surface adapted to a range of deflection angles can be obtained with rigid cylindrical rings or bands.

[13] C. S. Barrett and A. G. Guy, Rev. Sci. Inst. **15**, 13 (1944). E. Schwarz, Phys. Zeits. Sowietunion [5] **3**, 443 (1934).

[14] M. Straumanis and O. Mallis, Zeits. f. Physik **94**, 184 (1935) and M. Straumanis and A. Jevins, Zeits. f. Physik **98**, 461 (1936).

[15] Bradley, Lipson, and Patch, J. Sci. Inst. **18**, 216 (1941).

[16] M. J. Buerger, J. App. Phys. **16**, 501 (1945). See also *X-Ray Crystallography* (John Wiley and Sons, Inc., New York, 1942).

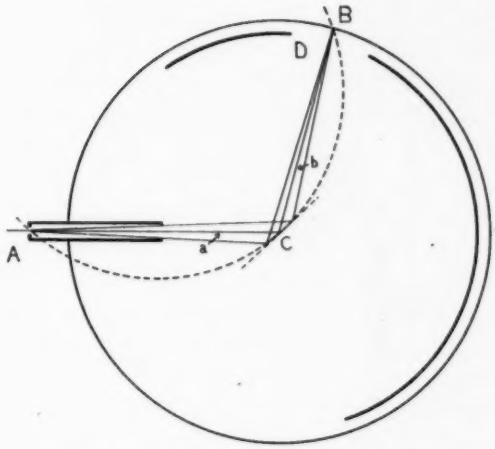


FIG. 6. Parafocusing central powder layer.

As a derivation from the toroidal surface the writer proposed an annular band.[17] The powder is placed on the inside surface C of a ring (Fig. 3) the axis of which comprises the points A and B. The radiation from A will then be deflected to B for a range of angles which can be varied by displacing the ring parallel to itself and displacing point B along the axis. B is represented by a small opening in a screen closely behind which the photographic film is placed. With the displacement of the ring C and of B the photographic film is moved at right angles to the axis AB. In praxis only part of the ring is used. The arrangement gives very high intensities. It does not, however, comply well with other demands when photographic recording is used because of the oblique incidence of the diffracted x-rays on the film, which is unfavorable for quantitative measurements. The arrangement may be useful in combination with a counter or ionization chamber.

A longitudinal strip is used in the arrangement of Seemann[18] and Bohlin[19] (Fig. 4) in which the toroidal surface is approximated by a cylindrical area tangential to a longitudinal band of the toroid. This arrangement is very simple in design and capable of recording a large angular range on one single film strip. Line distances are not proportionate to angular distances; it is unfavorable for intensity measurements and the meas-

[17] J. C. M. Brentano, Arch. Sci. Phys. et Nat. [5] **1**, 550 (1919).

[18] H. Seemann, Ann. d. Physik **59**, 455 (1919).

[19] H. Bohlin, Ann. d. Physik **61**, 421 (1920).

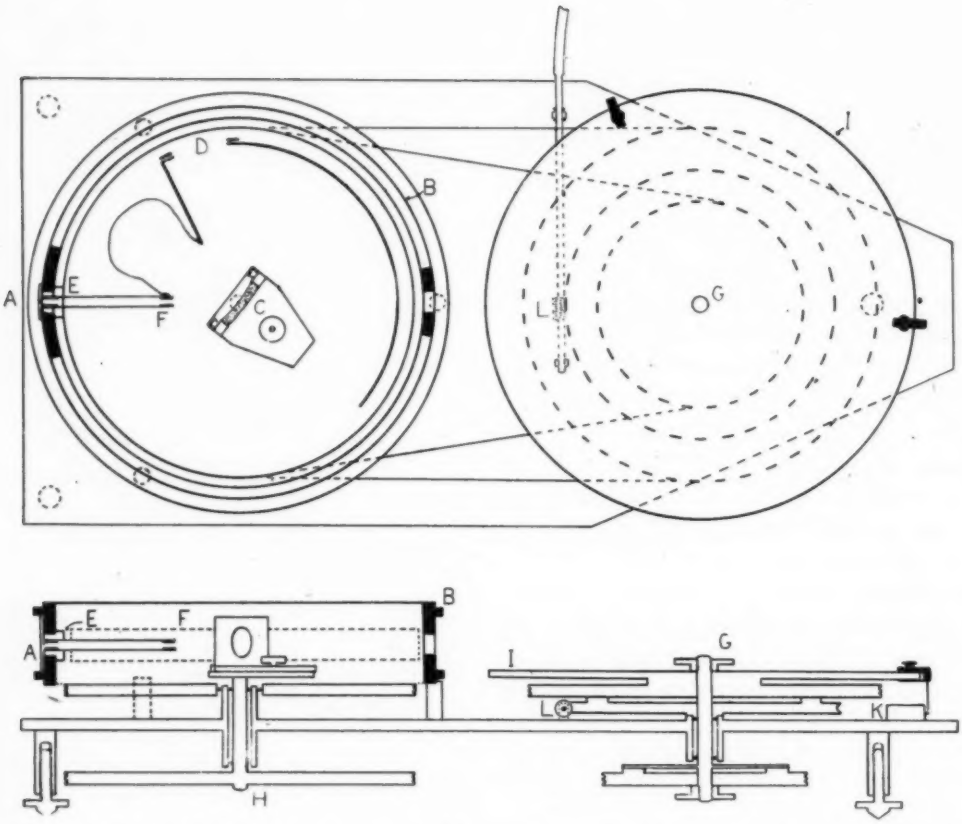


FIG. 7. Schematic diagram of camera with central parafocusing powder layer. The film is wrapped round the cylindrical camera body *B* and is exposed through slots. The collimating system consists of a slit in a front plate *A* and of inserts *E* and *F* in the collimating tube. Insert *E* determines the slit height; insert *F* determines the angular aperture. Slit and inserts are exchangeable. Aperture *D* can also be varied. A reversible motor through worm drive in *L* rotates shaft *G*. Other details are given in reference 24.

urement of line contours because the diffracted rays strike the photographic film with varying obliquity and because the entrance diaphragm in *A* is reproduced from the various parts of the powder surface at different rates of magnification. The method is, however, very well adapted for utilizing large volumes of powder in conjunction with incident beams of large aperture. Several asymmetrical cameras of this type have been described. De Jong[20] described a symmetrical camera, the cylindrical powder surface being opposite to the entrance slit. This is particularly suitable for back reflection records.

An arrangement, in which the photographic film forms a cylindrical surface while a thin powder layer forms a small area of the same surface (Fig. 5) and a converging incident beam is transmitted through the powder with the center of convergence situated on the cylindrical surface, has been proposed by Guinier[21] who makes use of a monochromator for obtaining the converging beam. Frevel[22] has proposed a similar arrangement using a converging Soller slit. While this arrangement can be discussed in terms of parafocusing surfaces, the most direct treatment results from considering that for the plane of incidence the condition of equal convergence of the incident and diffracted beams is satisfied for points of convergence situated on the cylindrical surface. The characteristics of this method are similar to those of the Bohlin-Seemann camera. Its particular merit is that very small diffraction angles can be observed while, on the other hand,

[20] P. de Jong, Physica **1**, 23 (1927).

[21] A. Guinier, Comptes rendus **204**, 1115 (1937).
[22] L. K. Frevel, Rev. Sci. Inst. **8**, 475 (1937).

the incident beam is transmitted through the powder so that the intensities depend on the absorption of the powder. The arrangement of Bohlin-Seemann, de Jong, and Guinier satisfies the parafocusing condition essentially with respect to one plane.

6. GONIOMETER WITH A CENTRAL POWDER LAYER

A goniometer proposed by the writer[23] uses a powder layer in the center of a camera which satisfies the parafocusing condition for points of a cylindrical surface. It is intended to conform with those requirements set out in Section 3 which are satisfied in the Debye-Hull camera and to provide in particular proportionality between line distances and angles and normal incidence of the diffracted x-ray beam on the photographic film. In addition it establishes such conditions that the relative intensities of the reflections from thick powder layers are independent of the particular absorption characteristics of the powder specimen and are actually the same as would be obtained from a small powder volume with negligible absorption. It uses larger powder volumes and wider angular apertures than the Debye-Hull camera and requires shorter exposure times although they are longer than with the arrangements of Section 5. A detailed discussion of the geometrical-optical relations and aberrations resulting with this method has been given in an earlier paper[24]; it will here be sufficient to indicate its characteristic features.

Considering in the first place the case of photographic recording, it will be seen (Fig. 6) that the toroidal surface intersects the surface of the photographic film. The point of intersection B determines the angle for which the parafocusing condition is satisfied for a powder layer in C so oriented as to conform with the particular parafocusing toroidal surface. The recording of an extended angular range requires then a screen with an opening D which in succession uncovers different parts of the film surface as the powder layer is rotated about an axis normal to the plane ACB. If a and b are the distances of the powder layer from the collimating slit A and the cylindrical surface B and α and β are the glancing angles of the incident and reflected beam with the plane tangential to the powder surface in C, the parafocusing condition is satisfied when

$$\sin\alpha/\sin\beta = a/b. \tag{3}$$

The displacements of the powder layer and of the screen opening D are so correlated as to conform with the constant $\sin\alpha/\sin\beta$ ratio of (3) for the whole angular range. It will be seen that this constant ratio introduced in the absorption term (2) secures that this term has the same values for all angles and that relative intensities are independent of the absorption characteristics of the powder.

The parafocusing powder surface is in a general way a surface of double curvature, which changes with the diffraction angle. A first approximation to it is a flat surface; a much better approximation adapted to a particular range is a surface of double curvature—data for radii of curvature of such surfaces are given in the second paper[24]—and a spherical surface, when an extended angular range is to be covered and the angular aperture of the incident beam is not too large. For $a=b$ for the 2θ range from 60° to 150° a spherical surface with a radius of curvature $3/4b$ will give considerably better definition than a flat layer, permitting both in the horizontal and in the vertical plane[25] an incident beam of an angular aperture of 2°–3°. This range can be extended by a V-shaped horizontal slit in front of the collimator near to the powder which moves with the powder layer. Such a slit can be used to restrict the vertical aperture for small angles; it is also helpful in reducing the strong intensities in that range. When using such a slit, allowance for it must be made in computing intensities.[26] In a more elaborate arrangement the vertical and horizontal aperture of the beam can be so controlled as to keep the solid angle of the incident beam constant and to restrict either the horizontal or vertical aperture in such a way as to obtain the smallest aberration. Such devices

[23] J. C. M. Brentano, Proc. Phys. Soc. **37**, 184 (1925).
[24] J. C. M. Brentano, Proc. Phys. Soc. **49**, 61 (1937).

[25] For brevity in the following discussion, the term *horizontal* is used for the plane ACB in which the reflections are observed and the term *vertical* for the plane which comprises the axis of rotation of the powder layer.
[26] For accurate intensity measurements we prefer in general to use a beam of constant vertical aperture for the whole range.

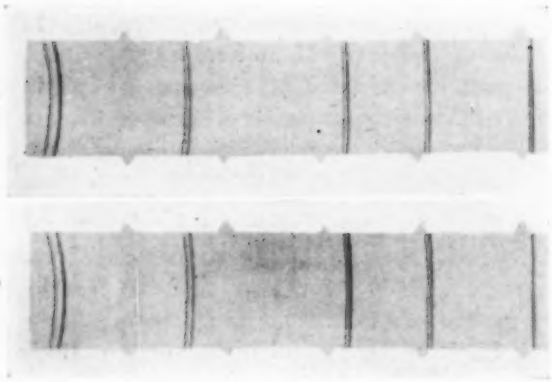


FIG. 8. Group of cadmium oxide lines recorded with parafocusing central powder layer (a) with rotating diaphragm, (b) without rotating diaphragm. Spherical layer: angular aperture of incident beam 2°.

are, however, of interest only then when it is desired to use beams of considerable angular aperture.

The schematic diagram (Fig. 7) indicates one design of the goniometer. The powder layer in C and a platform, on which the diaphragm D is mounted, rotate about a common axis H. They are driven by means of pulleys and steel wires from a shaft G. In combination with good ball bearings this system is very satisfactory. Projecting stops on a graduated disk I act on a microswitch K and determine the range of motion. For the ratio $a:b=1$, i.e., when the collimating slit A is on the cylindrical surface of the camera, the correct ratio of the angular displacements is 1:2. This is the simplest arrangement which can be satisfied with centric circular pulleys. The relative angular displacements corresponding to a ratio $a:b$ different from 1 can be obtained with pulleys of suitable contour. A spring loaded idler pulley is then required to maintain uniform tension. For correct intensities uniform velocity is given to the rotating screen while, when a and b differ, the angular speed of the powder layer varies. A ratio $a>b$ is used for obtaining very large apertures for the incident beam; slit A can then be moved into immediate vicinity of the target of the x-ray tube. This is discussed in Section 8.

The use of the rotating diaphragm in D implies that the exposure times increase with the angular range recorded. To compare the exposure times resulting from these conditions in recording a given range of say 90° with those of a Debye camera of equal size, it may be assumed that the outer pinhole of the latter and the entrance slit A of our camera are of equal area and that the aperture for the parafocusing beam is 2° by 2°, while the pinhole collimator with a length of 3″ has a pinhole diameter of 0.020″, as commonly used in some commercial instruments. The exposure required with the Debye camera is then 4.8 times greater than with the parafocusing central powder layer. This takes account of a diaphragm opening of 22° in D when the motion of diaphragm D exceeds by 22° on either side the 90° range, so as to secure uniform exposure for that range. When recording a group of lines comprised within a limited angular range, the rotating diaphragm can be dispensed with. The powder layer is then so set as to satisfy the parafocusing condition for the middle of this range. It can be rocked through a few degrees to avoid spotty lines. For the conditions considered above the intensity is then 29 times larger with the parafocusing layer than with the Debye camera. When a moderate amount of diffusion is permitted a range up to 60° can be recorded without using a moving diaphragm; this can be increased by reducing the angular aperture of the incident beam. Figure 8 shows a group of cadmium oxide lines recorded with and without diaphragm. The angular range covered is 65°.

When recording the diffraction pattern with a counter or ionization chamber its entrance slit is situated on circle B. The slit acts as diaphragm D and eliminates those aberrations which arise for angles away from the one for which the parafocusing relation is satisfied.[27] It does not eliminate the aberrations resulting from any deviation of the shape of the powder surface from the parafocusing toroid. The line broadening resulting from aberrations is then essentially determined by the deviations of shape and orientation of the actual surface from the toroid so that the use of correct surfaces and suitable diaphragms becomes important. When using ionization chambers and counters, which must be placed behind the circle B, the use of $a>b$ makes it possible to extend the angular range recorded without interference from the x-ray tube.

[27] Care must be used to avoid errors in evaluating intensities which result with too narrow slits.

We find it most convenient to carry out identifications on the ground of patterns covering a 2θ range extending from 60° to 150° and to disregard the ranges of very small and very large angles. For this middle range, particularly above 90°, the lines are well defined and any broadening can best be examined here. For the range of large and small angles the curvature of the lines, which results from the small aperture of the diffraction cones and which arises with all powder methods, is unfavorable for accurate measurements; besides the number of lines in these ranges is generally small.

For evaluating intensities the simplification with this arrangement results from the fact that by satisfying (3) for a fixed ratio a/b the absorption term (2) becomes independent of the diffraction angle as indicated above. The relative intensities are thus independent of the absorption coefficient of the material and are those for a small volume element with negligible absorption. They are thus directly comparable with those obtained from a highly diluted Debye powder rod. It should be noted that this independence from the deflection angles does hold good only when satisfying the parafocusing condition. When recording an extended angular range without the revolving diaphragm D the absorption becomes the function of the angular setting; it can be calculated for each line.[4]

The large angular apertures and the use of thick all-absorbing layers increase the powder volumes utilized and tend to suppress the spottiness caused by the presence of individual large crystallites. This can be further reduced by using a flat powder layer in form of a rotating disk with the axis of rotation off-set with respect to the incident beam. This disk takes the place of the powder layer in the center of the camera. The volume of powder and the angular range through which it is exposed to the incident x-ray beam is thus increased. Figure 9 shows the patterns of the same material taken with a rotated and displaced powder rod and with a rotated flat disk in the way described.

7. BACK REFLECTION CAMERAS

One particular type of goniometer is that for taking "back reflection" patterns, limited to large 2θ values. Changes of line position and of line width and contour in that range are most sensitive to changes of lattice spacings and to imperfections of the microcrystalline domains. In a widely used type of camera[28] the pattern is recorded on a flat film, a pinhole collimating tube being placed between the x-ray source and the recording film. By using a narrow beam similar to that used with the Debye-Hull camera sharp lines can be obtained over a range of angles without having to conform with parafocusing conditions. As with the Debye camera the exposure times are then long.

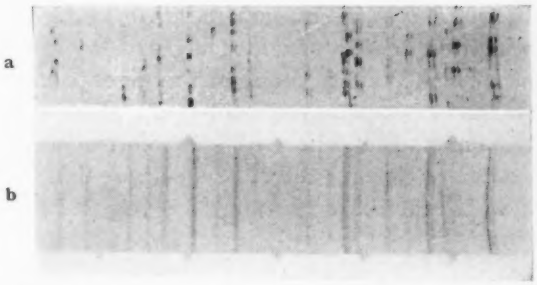


FIG. 9. Patterns of the same material taken (a) with a rotated and longitudinally displaced rod and (b) with the camera of Section 6 using a rotating disk.

By placing the outer collimating pinhole into a definite relation to the flat photographic film a beam of larger aperture can be used; good definition is then secured for one particular angle.[29] A flat or spherical powder layer in C (Fig. 10) is only an approximation to the parafocusing surface which is not spherical but has a radius of curvature $s/2$ in the horizontal plane and approximately s in the vertical plane, s being the distance from A to C. The angular aperture should thus not be too large. The merit of these arrangements is that with circular pinholes and diaphragms they possess azimuthal symmetry. They can thus be used for recording fiber structures and preferential orientations.

For a powder with random orientation a camera conforming with the toroid is preferable since it gives good definition for an extended angular range. By substituting for the cylindrical powder layer used by de Jong a spherical powder layer (Fig. 11) with a radius of curvature equal to

[28] G. Sachs and J. Weerts, Zeits. f. Physik **60**, 481 (1930).
[29] R. Berthold, Zeits f. Physik **15**, 42 (1934).

$3b/2$, where b is the camera radius, a better approximation to the parafocusing surface is obtained which takes account of the divergence of the beam in two planes and not in one plane only. Large apertures can be used with short exposure times except that for the 2θ range approximating 180° where the diffraction lines are very strongly

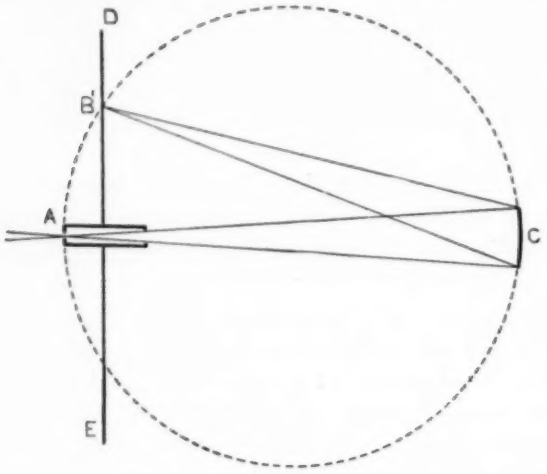


FIG. 10. Back reflection camera for recording preferential orientations. The pattern is projected on a flat photographic film in DE. The arc in C indicates the intersection with the parafocusing surface which, however, is not spherical but a surface of double curvature. When using a flat or spherical powder surface and circular pinholes for obtaining uniform conditions for all azimuthal angles, the aperture should thus not be too large.

curved, the vertical aperture should be somewhat restricted. By rotating this spherical surface about an axis somewhat offset to the direction of the incident beam the quality of the lines, when large crystallites are present, is greatly improved. Restricting the aperture of the beam in the horizontal plane will reduce the residual aberration in that direction and extend the angular range of good definition. A slit in A can be substituted for a pinhole as previously discussed in Section 6, except that its height should be limited for recording reflections for which 2θ approximates 180°; for such angles, however, the resolving power is so great that good definition can often be dispensed with. The line distances are not proportionate to the angles, a disadvantage which, however, is felt less for large angles since the higher resolving power makes it easier to dispense with accurate linear interpolations. The goniometer of Section 6 will record back reflection patterns and at the same time give proportionality between angular and linear distances. When used without rotating screen with the powder layer set at right angles to the incident beam symmetrical patterns are obtained; best definition is then satisfied for $2\theta = 180°$ where it is least relevant.

A more effective arrangement is obtained by placing A within the camera (Fig. 12) and setting the powder layer at such an angle as to be tangential to the parafocusing surface for a point B', which forms the center of the particular range for which good definition is required. It will be seen that the toroid (dotted circle) approximates the circle B on which the film is placed, so that for the projection on B a linear relationship between angles and film distances is secured, while at the same time approximating the parafocusing condition over an appreciable range. The radius of the spherical powder surface should be approximately $b/2$. Here again for the central strip of the photographic film the residual aberration in the vertical plane is not disturbing. This arrangement is very convenient for obtaining rapid data in that angular range of 150°–170°, which is most relevant for back reflection. While the ratio $\sin\alpha/\sin\beta$ is not constant, it varies but little with angle for large 2θ values so that the relative intensities approximate those obtained with the rotating screen.

8. COLLIMATING SYSTEMS

Three types of collimating systems are in use: The pinhole system, consisting of two pinholes placed at a distance which is large compared with their diameter, defining a narrow, approximately parallel beam. The Soller system, which gives beams with small angular divergence in one plane and large angular divergence in a plane at right angles. The system comprising a pinhole or slit which acts as a secondary source, and a second larger diaphragm which defines a beam comprising a definite solid angle.

The first is used in the Debye-Hull camera and in the usual back reflection camera, in which the location of the pinholes is irrelevant provided the collimator is sufficiently near to the powder to prevent an undue broadening of the beam. The Soller system, which in its angular spread is similar to that given by a Cauchois monochromator, meets the condition where focusing or

parafocusing is satisfied for one plane. It gives an angular spread similar to that of a Cauchois monochromator and Guinier and Frevel have used one and the other in the arrangement of Section 5. Beams comprising definite solid angles are used when the parafocusing condition is satisfied for two vertical planes. In comparing the merits of slits and of circular pinholes it should be considered that when the curvature of the diffraction lines is small, i.e., outside the range of very small or very large angles, some residual aberration in the direction of the lines is irrelevant. Slits oriented parallel to the direction of the lines can then be used without impairing the definition. The slit has then the advantage not only of giving greater intensity, but also of increasing the angular range through which the individual volume element is exposed to the incident x-ray beam. This improves the quality of the lines.

With any collimating system the intensity of the beam is independent of the distance of the entrance slit from the target of the x-ray tube so long as the beam, as defined by the slit, pinholes or diaphragms, projected backwards on the target covers an area which falls within the emitting anode spot. With the usual commercial tubes and pinhole diameters of 0.02″ this is satisfied when the distance between the entrance pinholes and tube target does not exceed the collimator length. There is then no gain in bringing the entrance slit nearer to the target. When conforming with parafocusing conditions with a collimator comprising a slit which acts as secondary source, the location of this slit with respect to the powder is relevant. At the same time with the much larger angular apertures used this requires bringing the slit or secondary source to the target; with tubes with small window to target distances apertures of 2°–3° can thus be obtained. When much larger apertures are required for obtaining extremely short exposures, the slit can be placed within the x-ray tube or the target used as slit. The writer has done this with a laboratory type tube operated from a pump[23] which in general design is similar to the new type Machlett tube. An alternative way of satisfying the requirements of large apertures is to use an extended anode spot. With an electron beam concentrated on a very small target area a moderate broadening of the spot does not effectively increase the loading capacity, because the volume of the anode surrounding the spot determines the rate at which heat can be dissipated; this is only little changed when the target area is increased by a small amount. With an extended spot, however, on a thin target and with forced water circulation the total loading capacity of the tube and the x-ray output can be considerably increased.

9. THE EVALUATION OF INTENSITIES

The conditions for quantitative evaluations are met most directly by the method of Section 6, which we developed primarily for the purpose of eliminating the effect of absorption on the relative intensities of the reflections recorded over an extended angular range. For identifications, so far as they depend on the relative line intensities, it is an obvious advantage if these are unaffected by the absorption in the particular powder specimen. For quantitative determinations on mixed powders[29] and the determination of unknown structure factors in terms of structure factors which are known, to which we have referred in the earlier part, reflections occurring for different angles can thus be correlated.

In such measurements it is essential to avoid preferential orientation of the powder particles. One way of achieving this is to place the powder with the binding material in a frame which in preparing the layer is held against a surface of the desired curvature, the powder being intro-

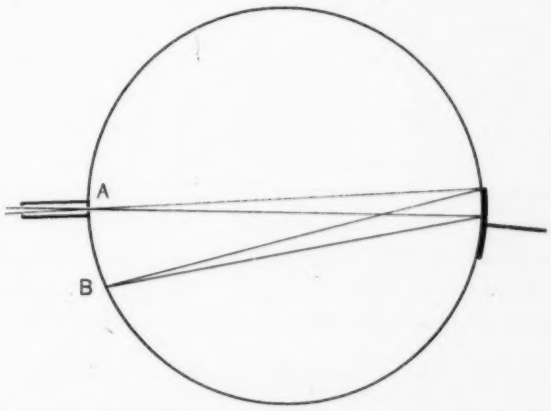


FIG. 11. Back reflection camera for recording powder diagrams on a cylindrical surface *B* using a spherical powder layer rotated about an offset axis.

duced from the back of the frame so as to avoid any sliding of the material near the front surface. It is then also required that the aperture of the incident beam, as defined by the collimator, should be fully utilized for the whole angular range and that the powder layer should be sufficiently thick to be practically all-absorbing. While the large volume of powder utilized is an advantage in improving the quality of the lines, this excludes the method in such instances where very little material is available. Eventually the condition of full absorption can be satisfied with materials constituted of light atoms by an admixture for which the absorption coefficient is high. When evaluating the relative intensities for mixed powders, the mixture should be sufficiently fine for the domains of heterogeneity to be small in terms of the distance for which appreciable absorption occurs.[30]

The goniometer with a central powder layer can be adapted to methods relating diffracted to incident intensities[31] and to the substitution method and to the evaluation of intensities from powder layers at fixed angles, discussed by Brindley and Spiers[32] and by the writer.[33] It can also readily be adapted to the recording of the diffraction lines on a rocking film, a method recently described by the writer[34] which considerably improves quantitative evaluations of lines which have an unfavorable line contour.

10. THE EVALUATION OF DIFFRACTION ANGLES

One way of correlating measured line positions and diffraction angles is to measure the distances of corresponding lines on a symmetrical pattern taken on both sides of a cylindrical camera. This is generally done in the evaluation of Debye patterns. In an alternative way fiducial marks are placed on the film. In a third method reference lines are recorded on the pattern by admixing to the powder a substance for which the spacings are well known. With all these methods interpolations are facilitated when the line distances are proportionate to angular distances.

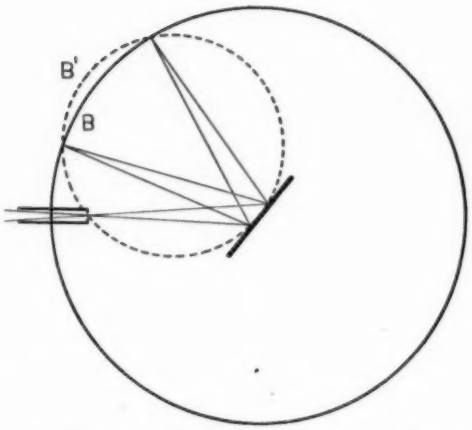


FIG. 12. Back reflection camera for recording powder photographs on a cylindrical surface B with line distances proportionate to angular distances.

By making use of symmetrical patterns in the technique of Straumanis,[14] developed for the Debye-Hull camera, errors due to faulty positioning of the specimen can be eliminated. This does not eliminate residual errors which arise from the penetration and absorption of x-rays in the powder rod and errors which result from uneven film shrinkage, errors which may be neglected for ordinary purposes but which become appreciable in view of the long distances over which the comparison measurements extend when high accuracy is demanded. Fiducial marks offer the simplest evaluation in terms of points on the circumference of the camera. The location of such marks represented by edges separating dark and light fields on the film is affected by the photographic Eberhard effect, a source of errors which has been fully discussed in relation to optical spectroscopy. This error is avoided with marks consisting of V-shaped dents or projections. These can be located to the same or to a higher degree of accuracy than the diffraction lines.[35] Such regularly spaced marks eliminate errors due to uneven film shrinkage; they do not in themselves, however, take care of errors due to faulty positioning of the powder.

Errors of asymmetry and eccentricity should be considered as arising from two different sources. One results from faults of workmanship;

[30] J. C. M. Brentano, Proc. Phys. Soc. **47**, 932 (1935).
[31] J. C. M. Brentano, Zeits. f. Physik **99**, 65 (1936).
[32] Brindley and Spiers, Proc. Leeds Phil. Lit. Soc. **3**, 4 and 73 (1935); Proc. Phys. Soc. **50**, 17 (1938).
[33] J. C. M. Brentano, Proc. Phys. Soc. **50**, 247 (1938).
[34] J. C. M. Brentano, J. Opt. Soc. Am. **35**, 382 (1945).

[35] A way of utilizing these marks in the microdensitometric evaluations is to trace with the edge of a razor blade fine scratches between the corresponding apexes of the V's.

in the camera with a parafocusing central layer such an error would be a deviation between the axis of rotation of the powder and the cylindrical surface on which the film is placed or faulty location of the entrance slit of the counter with respect to the graduated scale on which the angles are read. By careful designing and workmanship such errors can be reduced to limits smaller than the accuracy to which line positions can be evaluated. It is thus indicated to suppress such errors to the limits of accuracy required rather than to correct for them in each exposure. The other results from faults in mounting of the individual specimen and the penetration of x-rays within the powder. Powder surfaces can be located to a fair degree of accuracy; for instance in the parafocusing camera with central layer, the surface of powder specimens, made by the use of jigs or molds in the way described in the last section, can be located within 0.2 mm. With a camera of 6″ diameter the accuracy is then sufficient for ordinary identification purposes. When a higher degree of accuracy is required, faults of location and penetration can be eliminated for this particular method by using a flat thin layer on a support (plastic or glass) transparent to x-rays[23] in combination with a beam of limited angular aperture. Two exposures are then recorded on the same film for which the powder layer has been rotated by 180°, so that opposite sides of the layer face the x-ray beam. If the axis of rotation does not go through the powder layer, double lines result, the correct position being half-way between the two peaks. The line positions defined in this way are related to the fiducial marks. The technique of using thin layers is only needed when the highest accuracy is demanded.

The admixture of a substance with known spacings eliminates all errors resulting from faulty positioning and absorption and also all other aberrations. To distinguish between the two line patterns and detect superpositions it is then often necessary to take two patterns, one with the admixture and one without.

11. CAMERA SIZES

The effect of camera sizes on the accuracy of angular measurements has been discussed in an earlier paper,[24] to which reference may be made to avoid repetition. It may here be pointed out, however, that the marginal error attached to the use of finite angular apertures—which generally increases with the square of the aperture —is essentially a relative angular error; for given apertures it is thus independent of the diameter of the circle on which the pattern is projected. The same applies to the error resulting from uneven film shrinkage, which is proportional to the film length and in terms of angles independent of the camera size. Errors of this group have in common that their effect on the relative accuracy is independent of the camera size. Errors of a second group, resulting from the diffusion of x-rays in the photographic material, from the penetration in the powder, and essentially also from faulty location of the specimen, involve a certain marginal linear uncertainty in the location of points on the film surface. The effect of such errors on the angular accuracy decreases with the camera size. For best efficiency such conditions will be established as to obtain a balance in which no single error is preponderant. The statement sometimes made in regard to Debye-Hull cameras, that the exposure times increases with the third power of the camera diameter, is thus without definite significance without indicating whether the pinhole diameters and collimator lengths are changed in proportion to the camera diameter. If all dimensions are changed in proportion, the angular apertures of the beams are maintained and an increase of camera size satisfies essentially the purpose of reducing the effect of the second group of errors without much change of exposure times. If, on the other hand the same pinholes are used with a larger camera in which the length of the collimator tube has been increased in proportion to the camera diameter, some of the aberrations are very effectively decreased. Such a change leads, however, to rapidly increasing exposure times; applied to a "balanced" setting it leads to conditions in which some of the errors are unduly suppressed for best efficiency. An adjustment to the specific conditions is demanded in each case.

We generally chose the diameter best suited for our requirements and neglected to pay

attention to standard camera dimensions, since we found it not too difficult to have a scale made so adjusted to the camera diameter that line distances can be read directly in degrees. It is, however, possible to adjust slit openings and apertures so as to yield best efficiency for a given camera diameter. For microdensitometric evaluations the broader lines, which for a given angular resolution are obtained with a larger camera, are equally good or more favorable than the sharper lines with a small camera and we prefer thus large camera diameters in order to suppress aberrations of the second group. For visual observation the greater density gradients resulting with a small camera giving a close pattern of fine lines is a distinct advantage.

A Study of Oxide Cathodes by X-Ray Diffraction Methods

Part I. Methods, Conversion Studies, and Thermal Expansion Coefficients

A. EISENSTEIN*
Radiation Laboratory,† Massachusetts Institute of Technology, Cambridge, Massachusetts
(Received November 13, 1945)

Two methods are described for obtaining x-ray diffraction patterns of oxide-coated cathodes. One method is used in the study of the conversion process in forming the oxide cathode while the other method is utilized in measuring the thermal expansion coefficients of barium, strontium, and thorium oxides. The conversion of an equal molar barium-strontium carbonate solid solution, $(BaSr)CO_3$, involves (1) crystal growth in the carbonate, (2) decomposition to the mixed oxides, BaO and SrO, (3) formation of the oxide solid solution, (BaSr)O, and (4) crystal growth in the oxide. A similar sequence of events is observed in the conversion of a mixed carbonate, $BaCO_3+SrCO_3$. Crystal and particle size growth of carbonates and crystal growth of oxides are investigated and possible relationships are discussed.

I. INTRODUCTION

THE thermionic emission properties of oxide cathodes are directly related to certain chemical and physical structures; it is, therefore, essential for future experimental and theoretical studies that an accurate description of these structures be available. X-ray and electron diffraction studies have supplied information regarding the chemical nature and physical structure of active oxide coatings which is obtainable by no other means.

Burgers[1] showed that when a mechanical mixture of barium carbonate and strontium carbonate was decomposed in vacuum during the formation[2] of an oxide cathode, a simple conversion to the mixed oxides, BaO and SrO, took place. Further heating caused the formation of a true solid solution, (BaSr)O, at a rate which was temperature dependent. The lattice constant of such a solid solution was found to vary linearly with composition and continuously from 100 percent BaO to 100 percent SrO. Cathodes of (BaSr)O which were flashed at a high temperature, when examined by x-rays, showed asymmetrical diffraction lines which were interpreted as indicating a preferential loss of BaO from the surface.

Benjamin and Rooksby[3] found that a coating which was initially a carbonate solid solution, $(BaSr)CO_3$, decomposed to give an oxide solid solution, (BaSr)O. A plot of d.c. thermionic emission *versus* relative proportions of BaO and SrO present in the solid solution indicated a dependence such that the optimum emission was ob-

---

*Now at the Research Laboratory of Electronics, Massachusetts Institute of Technology.

† This paper is based on work done for the Office of Scientific Research and Development under contract OEMsr-262 with the Radiation Laboratory, Massachusetts Institute of Technology.

[1] W. G. Burgers, Zeits. f. Physik **80**, 352 (1933).

[2] See e.g., J. P. Blewett, J. App. Phys. **10**, 668 (1939).

[3] M. Benjamin and H. P. Rooksby, Phil. Mag. **15**, 810 (1933).

served from a composition of about 60 mole percent SrO – 40 mole percent BaO. A later investigation[4] pointed out the relationship between the physical state of the oxide and the thermionic emission: higher emissions were obtained when the oxide was homogeneous throughout than when it consisted of two or more phases. Benjamin, Huck, and Jenkins[5] used x-ray diffraction to supplement optical microscope and emission microscope data in a study of the physical state, crystal structure, and emission properties of coatings prepared under varied conditions.

One hope for the diffraction type experiment has been to check the theory[6] of electron emission which supposed a nearly monomolecular layer of "free" alkali metal on the external surface of the semi-conductor oxide. Because of the relatively low penetrating power of electrons as compared to x-rays, electron diffraction methods are best suited to the examination of surface layers. Gaertner,[7] using this method, confirmed the findings of Burgers: namely, that a cathode of bulk composition BaO and SrO has a surface of pure SrO, some 300 to 2000 atomic layers thick,[8] formed immediately following the conversion from the carbonates. Darbyshire[9] found similar surface layers on cathodes of bulk composition (BaSr)O and SrO; however, cathodes converted from $BaCO_3$ gave surface patterns of BaO. Diffraction patterns taken during the conversion of the carbonate solid solution showed lines of only $(BaSr)CO_3$ or (BaSr)O. If an adsorbed monomolecular surface layer of Ba were bound to underlying O lattice sites of SrO, it would be expected to give rise to a diffraction pattern differing from that of SrO only in relative line intensities. A comparison[10] of the SrO surface patterns from cathodes of (a) (BaSr)O and (b) SrO suggests the possible presence of an active layer in the first case. A more recent critical examination[8] of such patterns showed no evidence for the supposed active monomolecular layer. One should regard both results with a degree of skepticism since (1) although a layer of Ba at lattice sites corresponding to the Sr of SrO has a scattering power of roughly twice that of Sr, it is only one of the some 20 surface layers contributing to the total scattering, and (2) an adsorbed layer of randomly spaced Ba would contribute only to the diffuse scattering.

Part I of this paper will discuss methods which were found useful in the furthur study of oxide coated cathodes by x-ray methods. The application of these techniques to the study of chemical and physical changes during conversion will be presented as well as the application to measurements of the thermal expansion coefficients of the oxides.

II. METHODS

Since the oxides of barium and strontium are chemically unstable in an atmosphere containing carbon dioxide or water vapor, special techniques were required to permit diffraction patterns of the active cathode coatings to be taken without exposing the cathode to the atmosphere. Two methods were used: the first, which consisted essentially in guarding the cathode against atmospherics by a protective coating, is particularly applicable to the study of a great many different cathodes removed from vacuum tubes; the second, which employed a special x-ray diode, allows a detailed study of single cathodes as a function of processing and life. The first method may also be used to investigate changes which occur as a function of life if a separate cathode is used for each point studied.

Although a method has been reported[3] for removing activated oxide cathodes from special test diodes to thin-wall, evacuated glass capillaries for x-ray examination, this technique is not suited to the examination of large numbers of cathodes nor to cathodes processed in standard vacuum tubes. More easily adapted to this study is a method[8] consisting of opening the vacuum tube in an airtight glass-windowed manipulation box in which the atmosphere is maintained chemically inert by an adequate flow of carefully dried nitrogen.* Rubber gloves, attached to the

[4] M. Benjamin and H. P. Rooksby, Phil. Mag. **16**, 519 (1933).
[5] M. Benjamin, R. J. Huck, and R. O. Jenkins, Proc. Phys. Soc. **50**, 345 (1938).
[6] J. A. Becker, Phys. Rev. **34**, 1323 (1929).
[7] H. Gaertner, Phil. Mag. **19**, 82 (1935).
[8] H. Huber and S. Wagener, Zeits. f. tech. Physik **23**, 1 (1942).
[9] J. A. Darbyshire, Proc. Phys. Soc. **50**, 635 (1938).
[10] J. A. Darbyshire, Proc. Phys. Soc. **50**, 964 (1938).

* Airco "Dry Nitrogen," 99.6 percent pure, passed through a H_2SO_4 bubbler and two $CaSO_4$ drying towers.

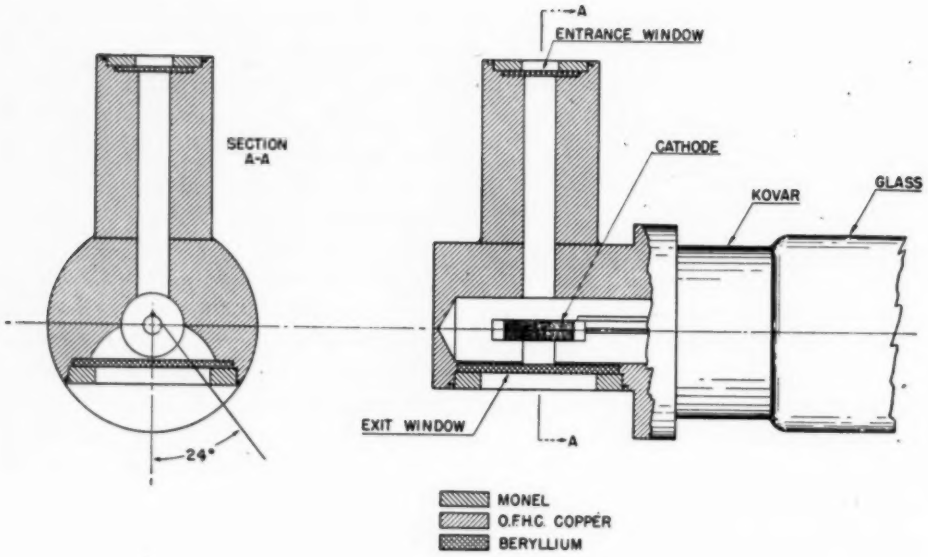


FIG. 1. Anode block for x-ray diode.

box by the wrist bands, permit this operation. After cracking the tube, the cathode is dipped in a protective coating, several of which were investigated. Polystyrene in a CCl_4 solvent gave adequate protection during periods of low humidity and was superior to nitrocellulose preparations. A special protective wax** coating was found to preserve the oxides without detectable chemical change for several days and in some instances for as long as three to four weeks. The use of the latter was generally adopted. Two weak diffraction lines are obtained from the wax corresponding to spacings of about 4.1A and 3.7A. Finally, the coated cathode was dipped in a suspension of tungsten powder in a nitrocellulose binder to provide a standard crystal size for calibration purposes.

Figure 1 shows the anode block of the special diode designed to study time changes occurring in a single cathode. The large beryllium*** exit window provides for the examination of diffracted rays up to 24° on either side of the incident beam. Primary radiation scattered from the entrance window is effectively checked by the built-in collimator tube, while scattering from the exit window is stopped by a strip of lead placed just outside the window. The windows, collimator, and Kovar ring were soldered to the block with a Ag-Cu eutectic alloy, without flux, in a H_2 atmosphere. Two diode types were used: in the first (see Figs. 1 and 2) the cathode is arranged to slide from its position between the windows to the upper part of the block where it is centered in a cylindrical anode section for operation and testing. In the second diode type, the cathode is mounted rigidly on the glass press which is separated from the anode block by a ground glass connection. This type is pumped continuously; it is not designed for emission studies but is used only to determine the effect of temperature and time on coating materials.

Since most of the cathodes which were subjected to diffraction analysis had an outside diameter of 3 mm, a 3-mm×0.5-mm collimator was used to provide a beam of nearly parallel x-rays incident over the semi-periphery of the cathode, Fig. 3A. Cameras of radius 4.70 cm were used in this study. As the sample was opaque to the characteristic radiation (filtered CuKα x-rays),* the diffraction pattern arose from conditions not generally utilized in x-ray analyses. Figure 3B shows a cathode of radius R, bathed in a parallel beam of radiation, scattering diffracted rays at an angle 2θ with the incident direction. Because of the very high absorption of the coating and base metal, only those portions of the periphery contained in S and S' can contribute

** A low softening temperature wax of the cerasin type.

*** Beryllium windows obtained from the Machlett Laboratories, Inc. Effective apertures and thicknesses, $\frac{1}{4}''\times.020''$ and $\frac{3}{4}''\times.040''$.

* Machlett type A-2 diffraction tube operated at 35 kv peak.

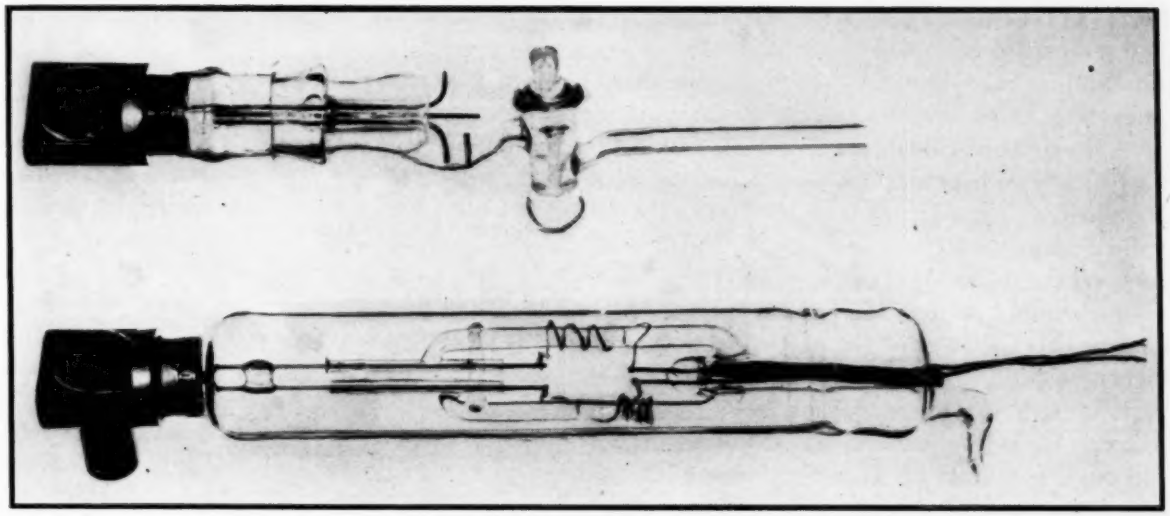


FIG. 2. Two types of x-ray diodes showing glass construction.

to the scattering AB and $A'B'$. For 2θ less than 90°, the diffracted rays are produced in two discrete areas so that inhomogeneities on the two sides of the irradiated cathode surface may be recognized by differences in the diffraction pattern on the two sides of the film.

Several disadvantages are encountered in the use of large diameter cathodes for x-ray studies which were avoided by previous workers[3] who used thin coated wires. The wide collimator necessary to completely bathe the sample may permit the use of a large target focal spot area and produce blurred diffraction lines. However, by setting the collimator slit perpendicular to the target face and adjusting for the radiation at a small angle, sharp lines are obtained. The large sample diameter introduces an appreciable error in the diffracted line position as recorded on the photographic film. This is a non-linear correction which depends critically upon the centering of the cathode within the camera as well as on the shape of the sample so that the usual correction[11] is not strictly applicable. From known diffraction lines recorded on each film from the protective wax, tungsten powder, and in some cases, the nickel base metal, a correction curve is obtained and applied to each unknown line position of the pattern.

III. CONVERSION STUDIES

That step in the processing of a vacuum tube in which the carbonate cathode coating is con-

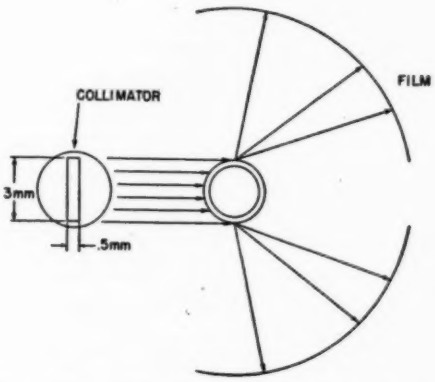


FIG. 3. (A). Slit collimator for bathing 3-mm cathode with incident beam.

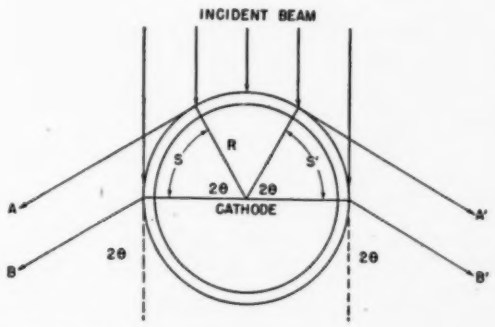


FIG. 3. (B). Showing effective scattering segments S and S' for opaque sample.

[11] R. Glocker, *Materialprüfung mit Röntgenstrahlen* (Verlagsbuchhandlung, Julius Springer, Berlin, 1936), p. 160.

verted to the oxide is generally considered critical and difficult to control. It is, indeed, this operation and its relationship to previous and subsequent treatment of the cathode materials that gives rise to the so-called "art of cathode making," particularly, when optimum thermionic emission is desired.

Benjamin, *et al.*,[5] reported the dependence of conversion rate and d.c. emission on the method of carbonate preparation which presumably allows variation in only the particle size. Thus, higher emission values were associated with small particle size carbonates which they believed convert to oxide particles of the same size. Although conversion and activation studies of mixed carbonates, $BaCO_3+SrCO_3$, are described[1,4] in some detail, a similar investigation of the conversion of the solid solution, $(BaSr)CO_3$, is not reported. Such a study should be of practical interest since solid solution carbonates are now used extensively in commercial vacuum tubes. In this report the conversion process shall refer to those steps in the processing of an oxide cathode which occur prior to the actual drawing of emission current and shall include, therefore, what has been referred to in the past as thermal activation of the oxide as well as decomposition of the carbonate.

TABLE I. Conversion schedule and heat treatment of $(BaSr)CO_3$ cathodes.

Time from previous step	Maximum temperature	Pressure	Pattern	Results
60 min.	450°C	10^{-7} mm	A	No physical or chemical change in carbonate coating
1	600	10^{-5}		No change
3	700	5×10^{-4}	B	Growth in carbonate crystal size
2	740	$>10^{-3}$	C	Crystal growth, SrO lines present
2	795	$>10^{-3}$	E–F	$BaCO_3$ and SrO on one side, BaO and SrO on opposite side
4	810	$<10^{-6}$	G	BaO and SrO
	Subsequent heating after decomposition			
30	800	10^{-7}	I	(BaSr)O
120	800	10^{-7}		(BaSr)O
5	850	10^{-7}		(BaSr)O
5	890	10^{-7}		(BaSr)O
5	940	10^{-7}		(BaSr)O
5	1000	10^{-7}		(BaSr)O, large crystals
5	1050	10^{-7}	J	(BaSr)O, very large crystals

The conversion of an equal molar barium-strontium solid solution carbonate,* $(BaSr)CO_3$, was studied by means of a series of x-ray diffraction patterns. Each of a series of identical cathodes** was partially converted and stopped at a different point in a standard conversion schedule. Each cathode was mounted in a glass bulb which was evacuated by means of a three-stage oil diffusion pump. Following a 60-minute bake-out at 450°C, cathodes were heated to about 700°C and held at that temperature until the carbon residue from the nitrocellulose coating binder was no longer visible. The heater current was then increased in regular steps, so chosen that complete decomposition would take place in about 15 minutes. Finally, each cathode was prepared for x-ray examination using the wax technique previously described.

Table I shows temperature and pressure values recorded at the corresponding times the conversion schedule was interrupted and represents, therefore, a complete cathode conversion. Seven cathodes were given additional heat treatments indicated in the lower section of Table I. Each time and temperature listed represents the total treatment following the above decomposition schedule. Temperatures listed are based on uncorrected pyrometer readings made at the center of the cathode from which point the x-ray diffraction patterns were also taken. Values below 750°C were obtained from an extrapolated heater wattage *vs.* temperature plot. An ionization gauge and d.c. amplifier with a maximum amplification of 1000 was used to measure the pressure in the vacuum system.

Seven of the x-ray diffraction patterns listed in Table I are reproduced in Fig. 4A, B, C, E, F, G, I, and J. Pattern *A* of the coating, following a 30-minute bake-out at 450°C, is identical with the original carbonate pattern except that scattering from the nitrocellulose binder is now missing. The small crystal size of the carbonate coating is reflected in the diffuse nature of the diffraction "lines." All strong "lines" with the

* Coating Type C51-2, obtained from the Raytheon Manufacturing Co., Newton, Massachusetts.

** The cathodes were coated approximately 10 mg/cm², 8-mm coated length, on 3×14 mm indirectly heated, Grade "A" nickel sleeves.

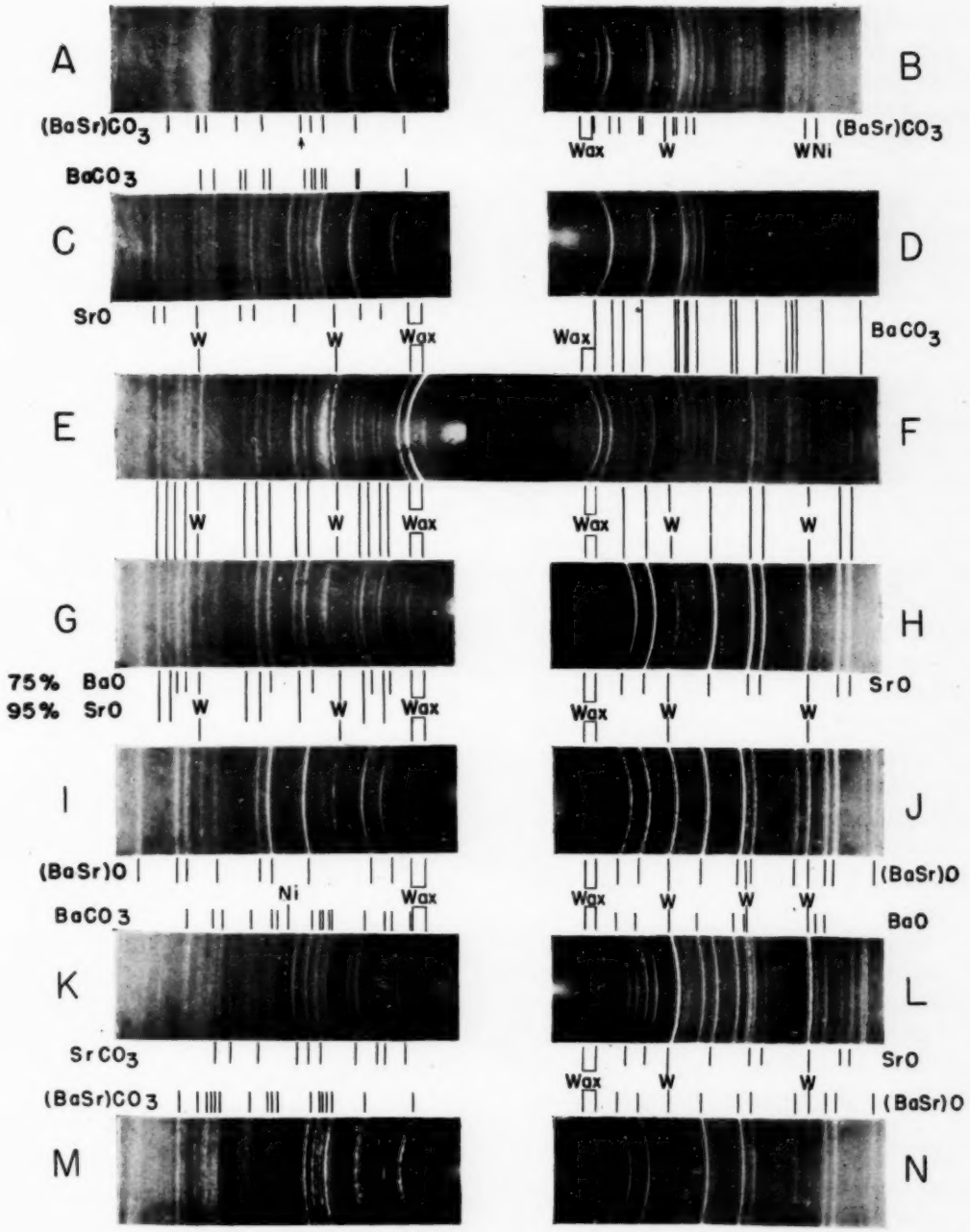


FIG. 4. X-ray diffraction patterns of cathode samples. See text.

exception of the one marked by an arrow are multiple diffraction lines. This may be seen by close inspection of *B* or *M* which are patterns of larger crystals of the same material. The origin of this multiple structure is not the presence of two or more solid solution phases since pure $BaCO_3$ in *D* shows a similar multiple line pattern. Attention is called to this point, for although the multiple nature of some of these lines is recognized,[12] they are frequently[13] treated as single reflections. Line breadth measurements for the determination of crystal size may thus be correctly applied only to the single line indicated. It is evident from the somewhat sharper lines in

[12] J. M. Cork and S. L. Gerhard, Am. Min. 16, 71 (1931).
[13] A.S.T.M. x-ray analysis card file.

pattern *B* of the cathode after 3 minutes at 700°C that crystal growth has begun although no chemical transformation has occurred. Pattern *C*, taken after an additional 2 minutes at 740°C, shows the beginning of the decomposition process. Lines of both $BaCO_3$ and SrO are present indicating the decomposition of $SrCO_3$ at this stage. Figure 4E and F shows an interesting diffraction pattern of a cathode at the 795°C stage. One side of the cathode gave lines only of $BaCO_3$ and SrO while the opposite side, which was probably at a slightly higher temperature, had completely converted to BaO and SrO. Solid solution formation had begun since two phases are present, 95 percent SrO – 5 percent BaO and 75 percent BaO – 25 percent SrO. (To aid in the identification of these compounds, compare with *D* of pure $BaCO_3$, *H* of pure SrO, and *L* of mixed crystals of BaO and SrO.) Inhomogeneities over the cathode surface frequently appear during this stage. Because of the extreme temperature sensitivity of these processes, composition variations are often encountered along the length of the coating. Pattern *G*, taken at the completion of the decomposition process at 810°C, also shows lines of the phases 95 percent SrO and 25 percent SrO. Subsequent heat treatment (Table I) shows that a homogeneous solid solution is formed in 30 minutes at 800°C, Pattern *I*, or in 5 minutes at 850°C.

For convenience in the further discussion of the conversion process a division into three separate stages will be made: (1) crystal growth in the carbonate prior to decomposition, (2) decomposition and solid solution formation, and, (3) crystal growth in the oxide following decomposition.

Decomposition and Solid Solution Formation

Although the presence of lines of both $BaCO_3$ and SrO in Fig. 4C and F shows the complete decomposition of the $SrCO_3$ component of the solid solution to SrO before the conversion of $BaCO_3$ reaches appreciable proportions, the exact mechanism of the process cannot be unambiguously described. Two mechanisms are suggested: (1) the dissociation of the carbonate solid solution into its separate components just prior to the conversion of $SrCO_3$, or (2) the conversion to an oxide solid solution and the immediate recarbonation of BaO in the conversion gasses.

Since the conversion products of an initial solid solution carbonate were mixed crystals of BaO and SrO, the conversion study was extended to mixed carbonates which were reported[1,4] to yield the same conversion products. Pattern 4-*K* is of a commercial mixed carbonate coating, RCA 33*C*-133. This coating, when decomposed at 800°C, using a conversion schedule similar to that for the solid solution carbonate, Table I, gave pattern 4-*L*. Lines of BaO and SrO mixed crystals are present. The rate of solid solution formation was investigated in a series of cathodes flashed, after decomposition, for 5 minutes at temperatures of 860°C, 900°C, 1000°C, 1050°C, and 1100°C. Only those patterns of cathodes flashed at 1000°C and above showed single lines characteristic of a solid solution. This relatively slow rate of solid solution formation is in agreement with the findings of others who report[1,4] that formation from mixed carbonates requires up to an hour at 925°C. Broad lines were found in these patterns which were interpreted as representing a range of solid solutions of near equal molar composition rather than as an indication of small crystal size.

It appears then that in the conversion of mixed ($BaCO_3$+$SrCO_3$) and solid solution (BaSr)CO_3 there is an increased rate of solid solution formation and a greater homogeneity in the final oxide solid solution in the case of (BaSr)CO_3. A comparison of *G* and *L*, both mixed oxides which were decomposition products of (BaSr)CO_3 and ($BaCO_3$+$SrCO_3$), respectively, showed no difference in crystal structure other than a possible variation in crystal size.

Oxide Crystal Size

Line breadth measurements made from microphotometer tracings were used in the conventional manner[14] to determine the average crystal size of carbonate and oxide samples. Corrections[15] for instrumental broadening effects were made from nearby large crystal tungsten lines. By means of this method the crystal size of both phase components of pattern *G* were found to lie between 135 and 180A, whereas, those of pattern *L* are in the range 240 to 325A. Further heating at

[14] M. v. Laue, Zeits. f. Krist. **64**, 115 (1926).
[15] B. E. Warren, J. Am. Ceram. Soc. **21**, 49 (1938).

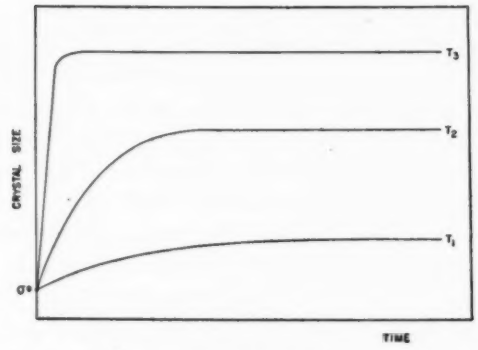


FIG. 5. Idealized crystal size growth curves of carbonates and oxides of initial size σ_0 growing at temperature $T_1 < T_2 < T_3$.

temperatures below 940°C did not result in appreciable crystal growth although solid solution formation did occur. Pattern *I* of the single oxide phase showed an average crystal size of 180A and after 2 hours at 800°C the size was 190A. Five minutes at 940°C gave a crystal size of 220A. Rapid growth occurs at 1050°C since the lines of pattern *J*, a cathode flashed for 5 minutes at this temperature, showed a decided "graininess" which is indicative of crystal sizes in excess of 10^{-3} cm. This suggests that the difference in crystal size of the mixed oxides, Fig. 4G and L, might be responsible for the observed differences in rate of formation and initial homogeneity of the final oxide solid solution. This seems to be a reasonable explanation since the formation of a homogeneous solid solution without crystal growth requires the interchange of at least half the molecules of any one crystal with adjacent crystals, a diffusion process which is expected to be strongly dependent on crystal size and on temperature in the manner indicated.

Carbonate Crystal Size

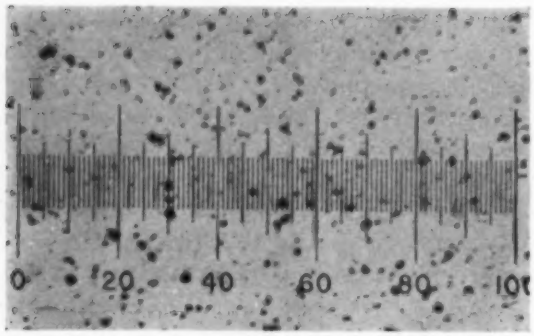
Since it is known that growth occurs in many crystals at temperatures above 400°C, and a correlation between emission and carbonate particle size is reported,[4] it was suspected that certain cathode conversion techniques might be associated with just such a particle growth to an optimum size. Three commercially prepared cathode coatings, (1) an equal molar $(BaSr)CO_3$,* (2) $BaCO_3$,** and (3) $SrCO_3$** were heat treated in a muffle furnace at temperatures of 450°C to 1050°C for times of from 5 minutes to 7 hours. These dry coatings were held in an enclosed nickel boat during treatment. X-ray diffraction patterns and photomicrographs were taken to measure both crystal and particle size of these samples. It should be noted that, although in some cases the crystal and particle sizes of a sample are identical, particles usually consist of a conglomerate of crystals, the size of which may be smaller by several orders of magnitude. Many of the samples heat treated above 600°C had crystal sizes in excess of 10^{-3} cm; hence, only semi-qualitative observations could be made based on the spot size[16] or on the degree of "graininess" of the diffraction lines. With heat treatment above 500°C, the crystal sizes of all three coatings were found to grow and to reach an equilibrium size which depended primarily on the temperature. This is a behavior which was observed in the growth of magnesium oxide[17] crystals as well as carbon black particles.[18] Idealized growth curves are shown in Fig. 5. Both the equilibrium size and growth rate are seen to increase with temperature. At 500°C all three coatings were found to increase from an initial crystal size of about 250A to greater than 10^{-3} cm in several hours, whereas at 600°C, a similar growth occurs in about 15 minutes. Still larger crystals were grown at higher temperatures by heating in an atmosphere of CO_2 to prevent decomposition.

The simultaneous growth of particle size may be seen in Fig. 6 which shows photomicrographs of $BaCO_3$ initially and after heating at 500°C for 15 minutes and 4 hours. A microscope magnification of 440× gives an over-all enlargement of 1.5 microns per division. Table II lists the minimum, average, and maximum size of the particles before and after treatment at 500°C and 600°C.

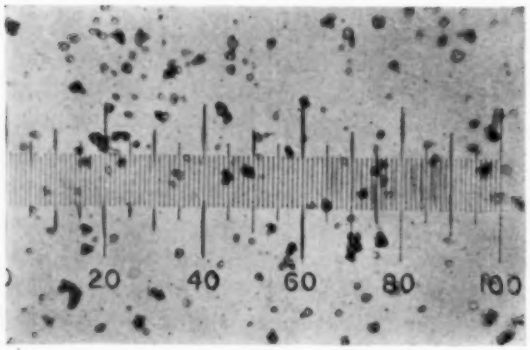
Thus, it seems that the average crystal size of untreated carbonates is in some cases less than 1/100 of the measured particle size; however, after growth (see Fig. 6), the particles appear to be single crystals. Particle growth at 600°C takes place about three times as rapidly as is observed

* Baker No. 3 Radio Mixture.
** Obtained from Raytheon Manufacturing Company.

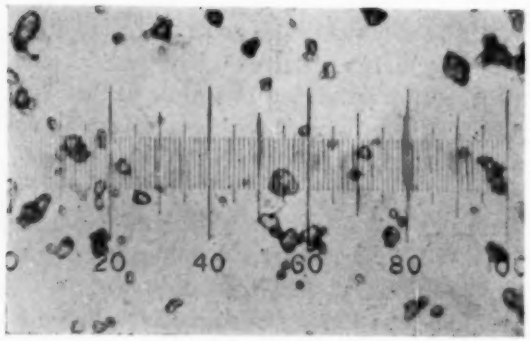
[16] G. L. Clark, *Applied X-Rays* (McGraw-Hill Book Company, Inc., New York, 1940), p. 506.
[17] J. A. Neuendorffer, M.I.T. Thesis (1941).
[18] J. Biscoe and B. E. Warren, J. App. Phys. **13**, 364 (1942).



(A)



(B)



(C)

FIG. 6. Growth in particle size of $BaCO_3$ induced by heat treatment. (A) Initial carbonate, (B) 15 min. at 500°C, (C) 4 hr. at 500°C. Magnification 1.5 microns per division.

at 500°C. Samples heated at 450°C showed only a slight increase in size.

Attempts made to influence the oxide crystal size by pretreatment of the carbonate or by variations in the conversion schedule were not successful. As with the carbonate, the crystal size of the oxide approached an equilibrium value which depended primarily on the temperature. In the temperature range of crystal growth this equilibrium size is reached in a matter of minutes so that the oxide crystal size can be expressed solely as a function of the maximum cathode temperature reached at any time during life. Cathodes converted at 800°C and flashed momentarily at 1000°C have crystals of a size characteristic of those converted at 1000°C, Fig. 4J.

A further indication of the unrelated nature of the carbonate and oxide crystal size is seen by comparing $(BaSr)CO_3$ heated to 800°C in CO_2 to prevent conversion and (BaSr)O converted from this large crystal carbonate at the same temperature in vacuum. The carbonate equilibrium size appears to be at least an order of magnitude larger than the equilibrium size of the oxide at the same temperature, Fig. 4M and N.

No evidence was found to indicate a change in crystal size during the normal life of an oxide cathode provided that the initial maximum cathode temperature is not exceeded.

It has been shown that (1) both the carbonate crystal size and particle size begin to increase at about 500°C and approach equilibrium values which depend on the time and temperature of treatment, (2) a similar crystal growth in the oxide solid solution begins at about 950°C, and (3) the oxide crystal size depends primarily on the maximum temperature of treatment above 950°C and bears no relationship to the crystal or particle size of the carbonate from which it was formed. It is probable, therefore, that no correlation exists between carbonate particle size and oxide particle size. Thus, it is difficult to see how the emission can depend upon the carbonate particle size in a fundamental manner.

IV. THERMAL EXPANSION COEFFICIENTS

The adherence of coating to base metal is an important parameter in the study of the thermionic emission of a variety of oxide coating, base metal combinations. For example, $BaCO_3$ on a chrome-plated nickel sleeve converted to BaO in vacuum at a relatively low temperature was found to fall from the sleeve on cooling, suggesting a possible difference in the coefficients of expansion of BaO and Cr. It was noted, however, that if this cathode was heated to a high temperature, the adherence was improved. A similarly poor bonding between base metal and coating has been observed in the case of ThO_2 on W.

Values of the thermal coefficients of expansion of most of the alkaline earth oxides not reported were determined in the following manner. The oxide or carbonate to be tested was sprayed on a cathode sleeve and mounted in a continuously pumped, demountable x-ray diode (Fig. 2). The diode was rigidly mounted with respect to the x-ray tube, collimator, and camera and sealed to a high vacuum system. With this arrangement one x-ray diffraction pattern was taken of the oxide at room temperature and a second at some elevated temperature, usually 875°C. The fractional change of lattice constant $\Delta d/d$ with temperature ΔT is simply[19] the coefficient of expansion for crystals having a cubic structure. With this method the lattice constant can be measured to an accuracy of 0.2 percent and since the change in this parameter is about 1.5 percent at 875°C for BaO, an over-all accuracy of about 15 percent is indicated.

TABLE II. Particle growth in carbonate at 500°C and 600°C.

	$BaCO_3$			$SrCO_3$			$(BaSr)CO_3$		
Time	Min.	Av.	Max.	Min.	Av.	Max.	Min.	Av.	Max.
0 min.	0.5μ	2μ	4μ	0.5μ	2μ	4μ	1μ	3μ	9μ
				500°C					
15 min.	1	3	6	1	3	5	1.5	3	6
4 hr.	1.5	4	8	2	5	11	2	5	10
				600°C					
5 min.	1.5	2.5	7	1	3	5	1.5	5	12
15 min.	1.8	3	7	1.5	4.5	12	2	9	20

Values of the linear coefficient of thermal expansion thus obtained for the various coatings tested are shown in Table III. Reported values for metals which are used for base metals are also shown in this table. Two factors should be considered in attempting to correlate expansion coefficients and bonding properties: (1) the values given in Table III described the true crystal expansion and represent, therefore, the maximum value that would be encountered in a low density sprayed coating, and (2) if an interface layer of different composition[20] is present between the coating and base metal, the adherence will depend on the interface coefficient as well. It is interesting to note, however, that the oxides have larger coefficients than their respective base metals so there is a gripping action of the coating on cooling.

TABLE III. Linear expansion coefficients of oxide coatings and base metals.

Material	$\Delta d/d\Delta T \times 10^6$
SrO	32.5±2.6
(BaSr)O	26.6±3.0
BaO	17.8±2.1
ThO_2	11.3±0.4
Nickel	13.5
Platinum	9.0
Copper	16.1
Tungsten	4.6

No structural change other than that caused by the lattice expansion was observed in the patterns of the coatings at 875°C.

The author is indebted to Dr. E. A. Coomes and members of the Cathode Research Group for valuable suggestions and assistance during the course of this investigation. The photomicrographs in Fig. 6 were kindly prepared by Mr. C. D. Prater of the Bartol Research Foundation.

[19] A. H. Compton and S. K. Allison, *X-Rays in Theory and Experiment* (D. Van Nostrand Co., Inc., New York, 1935), p. 682.

[20] H. D. Arnold, Phys. Rev. **16**, 70 (1920); A. Gehrts, Zeits. f. tech. Physik **11**, 246 (1930).

Methods for Betatron or Synchrotron Beam Removal

EUGENE C. CRITTENDEN, JR.* AND WILLIAM E. PARKINS
Radiation Laboratory, University of California, Berkeley, California
(Received January 9, 1946)

Two methods for bringing out a partially collimated electron beam from a betatron or synchrotron are discussed. Both of the methods become easier as the size of the accelerator is increased. One method makes use of a perturbing magnetic field to focus the electrons as they are made to leave the field of the accelerator by means of orbit expansion. The other method makes use of a pulsed deflecting system where the deflecting field is applied during a time short compared to the period of revolution of the electrons.

TO date devices of the betatron and synchrotron[1,2] type for accelerating electrons to high energies have not been adapted to bringing the electrons out in a beam for study, but rather have made use of the x-rays produced by the electrons striking some internal target. In the following, two methods are described which may be used to eject a partially collimated electron beam from such devices. These methods are made practical by the increased size of the present largest instruments in use or under construction. No experimental test has been made of the proposals, but the current importance of the beam removal problem has prompted submitting them without awaiting construction of a suitable apparatus at the Radiation Laboratory. The methods are suggested in connection with the removal of an electron beam although they may be equally useful when applied to ion beams, as in a frequency modulated cyclotron. They may be employed to advantage in any accelerating device using a magnetic field to restrain charged particles to move in cyclic paths such that the radial displacement from one turn to the next is normally small.

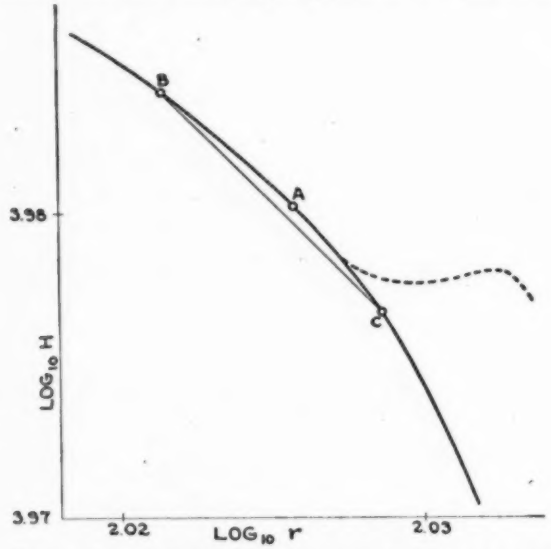


FIG. 1. Magnetic field to be expected at the edge of a synchrotron magnet of 1-meter orbit radius, based on data from a ¼-scale direct current model magnet. The dotted line represents the field as perturbed by the coil of Fig. 3a with a current of 510 amperes.

The magnetic field at any instant of time is usually expressed as:

$$H = H_0(r_0/r)^n,$$

where r is the distance from the center of the system, H_0 is the magnetic field at some fixed distance r_0 from the center of the system, and n is a dimensionless constant. At the edges of the pole face the field varies in a manner to which this form is not well suited since H_0, r_0, and n must all be considered variable. However n, represented as the negative of the slope of a log H *versus* log r plot, remains a useful quantity.

As the electron orbits are slowly expanded, as in the betatron by means of expanding coils, or in the synchrotron by means of turning off the oscillator at a time after the occurrence of peak magnetic field, the electrons will ultimately reach a value of r at which their orbits become unstable. The solid curve in Fig. 1 is typical of the field near the pole face edge in the median plane of the gap. Here instability occurs at point A, where $n=1$, if the electron has zero radial oscillation amplitude about its instantaneous orbit. If the electron has an appreciable radial oscillation

* On leave from Case School of Applied Science, Cleveland, Ohio.

[1] E. M. McMillan, Phys. Rev. **68**, 143 (1945).

[2] V. Veksler, J. Phys. U.S.S.R. **9**, 153 (1945).

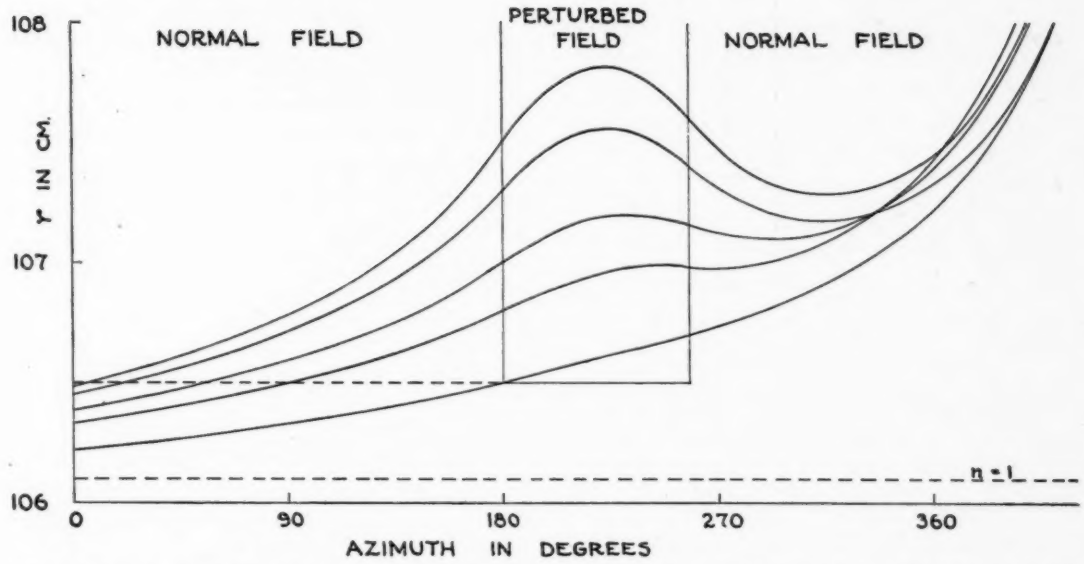


FIG. 2. Predicted electron trajectories for the magnetic field of Fig. 1 with a perturbed field given by the dotted line of Fig. 1 over 78° of azimuth.

amplitude, it will be released somewhat earlier in time. This condition is illustrated in Fig. 1 by an electron which has its instantaneous orbit at B, but by virtue of radial oscillation reaches a point of instability, C, determined by the intersection of the magnetic field curve and a straight line of slope -1 passing through B. After release the electrons move outward with increasing radial velocity. If allowed to continue, they leave the magnetic field of the accelerator producing a sheet of energetic electrons spread uniformly through 360° of azimuth.

The shape of magnetic field ordinarily encountered in a betatron or synchrotron is such that after an electron has passed a value of r about $\frac{1}{2}$ percent greater than r at the edge of instability, its value of r increases another one percent in traveling about 180° of azimuth. In a large accelerator this radial velocity is a convenient one for handling the released electrons in a focusing field. If the magnetic field is altered over a small range of azimuth in a manner similar to that represented by the dotted line in Fig. 1, the electrons can be made to emerge as a partially collimated beam. The electrons could be similarly handled in an electric field but proper shaping of an electric field seems to have more practical difficulties.

To illustrate a focusing scheme, results of calculations on one arrangement planned for a particular application will be described here. A perturbed field corresponding to the dotted line of Fig. 1 was employed over a range of 78° in azimuth. The electron trajectories were determined by the graphical method described in the paper immediately following. Figure 2 shows the trajectories for electrons of zero radial oscillation amplitude. The unperturbed magnetic field employed was that observed in a $\frac{1}{8}$-scale model of the magnet for a synchrotron of 1-meter radius. The perturbing field producing the focusing effects was the calculated field produced by the array of conductors shown in Fig. 3a where each conductor carries 510 amperes. The wires were arranged in a plane to permit construction by firing silver conductors into grooves in flat ceramic plates, thus assuring good mechanical support. The particular array was chosen to give negligible perturbing field inside the region of stability and to have a field profile giving small aberration, i.e., focusing effects too strong or too weak for electrons entering the perturbed region at the average radial displacement. This array of conductors gives a usable field depth of 1.0 cm in the axial direction. To avoid excessive values of the average power dissipation the coil need only be pulsed for a few microseconds when the electrons are expanding over the edge of instability. The use of local iron to produce the focusing field would avoid the complication of an

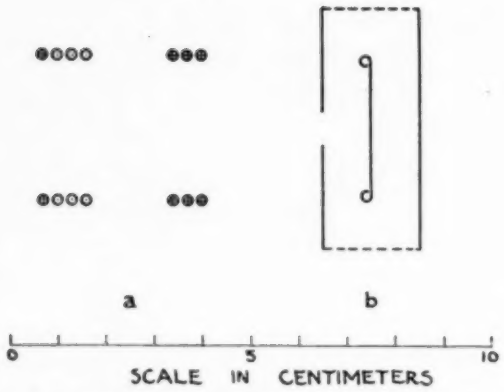


FIG. 3. (a) Cross section of the conductor array used to produce the perturbing field. Numbering from left to right, wires 1, 5, 6, and 7 carry current into the page, wires 2, 3, and 4 carry current out of the page. (b) Cross section of a deflecting electrode. The center of the magnet lies to the left for both diagrams.

electric pulsing circuit but has difficulties with regard to mechanical support, production of a correctly shaped field, and flexibility of adjustment. The axial defocusing inherent in the region of negative n occurring over a part of the dotted line of Fig. 1 is not serious because of the small path length of an electron in this field.

Figure 2 indicates that all trajectories originating in a range of azimuth of 174° would be brought into a beam crossing $r = 108.0$ cm within a range of azimuth of 16°. This beam would consequently contain 48 percent of the accelerated electrons. Tracing these rays out to where the field is zero predicts a beam with an angular divergence of 17°. In this example the inner edge of the region of perturbed field was placed at $r = 106.5$ cm whereas $n = 1$ at $r = 106.1$ cm. This was done to avoid introducing a perturbation in the orbits of radially oscillating electrons before they reached their edge of instability. Electrons which have their edge of instability at $r = 106.5$ cm would have a radial oscillation amplitude corresponding to a beam 1.5 cm wide in a region where $n = \frac{3}{4}$. These electrons are properly focused by the same coil but a current of 440 amperes is required in place of the 510 amperes for non-oscillating electrons. Since radially oscillating electrons are released early and at a time dependent on their amplitude, a properly varying current in the focusing coil could serve for all electrons.

If the maximum amplitude of radial oscillation proves to be less than the value assumed, which seems likely, the edge of the perturbing field could be moved closer to the circle at which $n = 1$. This would increase the fraction of the accelerated electrons appearing in the beam. An upper limit to the improvement attainable would be reached when the electrons began to pass through the outer edge of the perturbing field on the second revolution.

A modification which has some possibility for increasing the fraction of the electrons appearing in the beam, but that is difficult to evaluate without actual trial, is the application of a field variation in the region of the expanding orbits, over a small azimuthal range. The disturbance produced in the orbital motion during expansion might be such as to cause the electrons to exhibit a preferential range of azimuth in which they would cross their edge of instability. If this range were so located as to lie within the range from which electrons are successfully focused, an increase in the fraction of the electrons in the beam would result.

The second method to be discussed for producing a collimated beam of electrons is in effect an extreme case of concentrating the range of azimuth in which electrons pass into the region of instability. For large accelerators it seems probable this method could produce a fairly well collimated beam without the help of an auxiliary focusing field. By allowing the orbits to expand almost to their edge of instability, deflection of the entire beam could be accomplished with a reasonable electric or magnetic field, provided this field could be applied in the form of a flat topped pulse, rising in a time short compared to the period of the electrons. The fraction of the total number of electrons brought into a collimated beam would depend on the degree to which the electrical problems could be solved.

As an example, one possible solution would be the use of a charged transmission line of electrical length at least half the circumference of the accelerator, connected by means of spark discharge to an electric deflecting plate. The deflecting plate should resemble as nearly as possible a short continuation of the transmission line and be connected at its extreme end to a terminating resistance equal to the characteristic impedance of the transmission line. The time of firing of the spark is not critical and breakdown times of the

order of 10^{-9} second are known to be possible. If employed on an accelerator having a 1-meter orbit radius, where the period of the electrons is 2×10^{-8} second, this method should yield a reasonable fraction of the electrons in a collimated beam.

To illustrate the degree of collimation produced by such a deflector, consider the electrode structure shown in Fig. 3b. Such an arrangement should be constructed to spiral outward at the rate required to keep the deflected beam nearly grazing the inner electrode. With +80 kv applied to the transmission line, approximately +40 kv would appear on the inner electrode, and would give a fairly uniform electric field of 37,000 volts per cm in the region just to the left of the inner electrode. If the electrode is so connected that the conventional current flows in the direction opposite to the direction of motion of the electrons to be deflected, the magnetic field due to the current assists the deflection and is quite uniform in the region occupied by the beam. In the geometry considered here the force on a beam electron due to the magnetic field produced by current in the deflector will be almost 40 percent of the total deflecting force. For the deflecting electrode occupying 40° of azimuth and in the unperturbed magnetic field given by the solid line of Fig. 1, suppose the deflecting potential is applied at a time when the electron beam has expanded until its center lies 0.6 cm inside the circle at which $n=1$. In this case, considering both electric and magnetic fields produced by the deflector, graphical ray tracing predicts an angular divergence of 18° in the emergent beam for an internal beam width of 0.5 cm at the deflector. This corresponds to a beam 0.4 cm wide in an $n=\frac{3}{4}$ region. The angular divergence decreases approximately linearly with internal beam width and decreases with increasing applied voltage.

The foregoing has pointed out how electron beams might be withdrawn from large accelerators of the betatron or synchrotron type. One scheme uses a magnetic lens to concentrate electrons leaving with 360° spread in azimuth, the other employs a deflector which starts the outward motion in a small range of azimuth. At present the lens appears to be a slightly more effective and practical solution. If beam intensity and collimation were prime considerations, it is conceivable a combination of deflector and lens would comprise the best arrangement. These methods are not restricted to use with an electron beam but may be valuable when employed with an ion beam, such as is produced in a cyclotron.

The authors wish to acknowledge the information and advice given in connection with this problem by Professor E. M. McMillan.

A Graphical Method for Determining Particle Trajectories

WILLIAM E. PARKINS AND EUGENE C. CRITTENDEN, JR.
Radiation Laboratory, University of California, Berkeley, California

(Received January 9, 1946)

A graphical ray tracing scheme is described which gives improved accuracy for tracing particle orbits of the type encountered in the cyclotron, betatron, and synchrotron. The method allows a large increase in scale factor by elimination of the necessity of a compass arm. The motion of the particle is deduced from the motion of a working point, related to the instantaneous center of curvature of the particle orbit.

AT present there is an increasing number of applications for controlling the motion of charged particles by electric and magnetic fields. Except in simple cases the problem of determining trajectories is difficult. A graphical method is described here which is applicable in instances where the motion is essentially confined to a plane and where the radius of curvature of the particle is known as a function of space coordinates. In particular the method is useful for problems of motion of charged particles in devices similar to the cyclotron and betatron.

A graphical method in common use employs step-wise construction of the trajectory. A compass is set for a radius corresponding to the initial radius of curvature of path and placed on a layout of the coordinate system so that the compass arm extends normal to the initial path direction. Then the compass is swung through some angle which would bring the particle to a coordinate position where its radius of curvature is slightly changed. Here the compass arm length is changed correspondingly by moving its center without changing the arm position. Tracing may be continued by repeating the process of drawing a small arc and adjusting to a new radius of curvature. By taking sufficiently small steps any desired accuracy may be obtained within the mechanical limits of error of the graphical technique. But even this limitation may be reduced by using larger scale drawing instruments, and the method only breaks down when the accuracy required makes it too cumbersome to be practical.

In order to eliminate the cumbersome feature, the compass arm, a new system has been devised which essentially deduces the trajectory from the motion of the center of curvature of the particle. In cases where the center of curvature and origin remain within a region of dimensions small compared to the radii of curvature, any plot of the motion of the center of curvature can be done to a much larger scale than the particle trajectory. Qualitatively the method depends on determination of the progress of the particle in a polar coordinate system from the motion of a "working" point which is related to the center of curvature and rotated about the origin.

In practice the method can be carried out as follows. A piece of tracing paper is placed over a piece of graph paper serving as a rectangular coordinate system. A pin is stuck through some point designated as origin and into a drawing board below permitting rotation of the tracing paper with respect to the graph paper about this center. A working point is plotted on the tracing paper with Cartesian coordinates, $x_1=r_1 \sin \alpha_1$ and $y_1=r_1-R_1$, where r_1, R_1, and α_1 are known from the initial conditions of the particle and the properties of the field and represent, respectively, the polar coordinate radius of the particle, its radius of curvature, and the angle from the polar coordinate radius to the line connecting the particle and its center of curvature. A reference mark is also made to indicate the initial angular position of the tracing paper. Assuming the particle is traveling in an r, θ coordinate system, but the working point in a Cartesian system, some convenient increment of r is chosen and the tracing paper rotated until the working point moves a y distance equal to the chosen increment $\Delta_1 r$. The proper rotation of the tracing paper is in the direction opposite to the direction of rotation of the particle. It is as though the polar system in which the particle moves rotates at the proper rate to keep the particle on the positive y axis of the Cartesian system, while one origin is common to both systems. At this new setting an angular reference mark is made to determine the change in θ. The particle can now be assumed to be at $r_1+\Delta_1 r$ and $\theta_1+\Delta_1\phi$ where $\Delta_1\phi$ was the angular rotation of the tracing paper. For this new location the radius of curvature, R_2, is computed from a knowledge of field conditions. Without moving the tracing paper a new working point is plotted with y coordinate equal to r_2-R_2 where $r_2=r_1+\Delta_1 r$, and x coordinate the same as the new x value possessed by the first working point. The paper is rotated again to change the y coordinate of the second working point by $\Delta_2 r$ and an angular reference mark made. It is advisable to keep the increments of r some constant and convenient value as much as possible and to number successive angular reference marks together with the working points. This process is repeated while the progress of the particle in r and θ is tabulated, to be plotted later if desired. Figure 1 shows the type of pattern which would be produced if the path of the working point were projected through the tracing paper onto the coordinate paper.

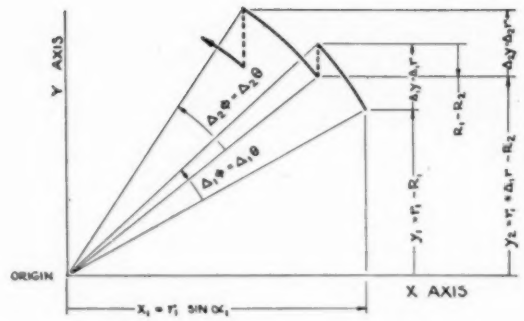


FIG. 1. Typical path of the working point shown projected through the tracing paper onto the coordinate paper.

If the field is a function of θ instead of r, or of both coordinates, it may be more convenient to use fixed increments of θ. Should the field be a function of θ alone, values of R computed for each increment of θ can serve to trace a variety of particles with different initial conditions. For a field which is a function of r alone, a single table of $r-R$ values can be used for different initial conditions. The value of R computed should be for some point within the interval concerned. Half-way points are sufficiently accurate except when the particle is passing a maximum within the interval. Here interpolation is recommended.

Consideration of the errors involved leads to the following. If the working point had been located at the center of curvature, its coordinates would have been:

$$x=R\sin\alpha,$$
$$y=r-R\cos\alpha.$$

To introduce no errors, the change in r, θ, and α of the particle must be exactly determined by the change in x and y of the working point during rotation of the tracing paper. It can be shown this is true here for a sufficiently small rotation only if

$$d\theta=d\phi,$$
$$dr=(r/R\cos\alpha)dy.$$

Hence, use of the center of curvature as the working point requires that Δr be calculated from Δy by use of r, R, and α. In cases where R/r is nearly one and α nearly zero this correction is hardly worth while. A more important consideration in the choice of working point coordinates is the difficulty of calculating them for each point plotted. Since r and α remain constant and R changes each time a new working point is plotted, use of the equations for the center of curvature to locate the working point proves inconvenient.

Suppose the coordinates of the working point were

$$x=r\sin\alpha,$$
$$y=r\cos\alpha-R.$$

Here an exact determination of r, θ, and α can be shown to require

$$d\theta=(R/r)\cos\alpha d\phi,$$
$$dr=(R/r)dy.$$

This choice has the disadvantage that $\Delta\theta$ must be calculated from $\Delta\phi$ and Δr from Δy, unless R/r is sufficiently near one and α sufficiently near zero. However a more than compensating feature lies in the fact that the x coordinate of each new working point is identical to the final x coordinate of the preceding point and the y coordinate is changed only by ΔR, the difference in R for the two intervals. If a procedure is desired which is exact regardless of the value of R/r or α, this last choice of working point coordinates is recommended. Use of sufficiently small intervals in plotting will determine a trajectory subject only to errors in graphical technique.

In the detailed description of the method the working point coordinates were given as

$$x=r\sin\alpha,$$
$$y=r-R.$$

The errors arising from the simplified use outlined for this choice of coordinates can be judged by comparison to the exact procedure just mentioned. The differences are $r-R$ replaces $r\cos\alpha-R$ for the y coordinate of the working point, $d\theta$ is assumed equal to $d\phi$ thus neglecting the $(R/r)\cos\alpha$ factor, and dr is assumed equal to dy neglecting the R/r factor. These differences all become negligible as R/r approaches one and α approaches zero. Errors of course are cumulative but the net error is usually reduced by R/r assuming values both greater than and less than one.

The advantage of the method discussed over the usual graphical method lies in the large increase in scale factor which can be employed in many instances. Such increased scale factors are possible when the center of curvature for the trajectory of the particle remains within a region of dimensions small compared to the radii of curvature. It is also true that under these conditions α will be small and R/r near unity contributing to the accuracy of the more rapid approximate scheme for locating and using the working points. This type of trajectory is characteristic of particle motion in the cyclotron, betatron, and synchrotron. Cases are not uncommon in which the motion permits a scale factor corresponding to an effective compass arm length of the order of 100 meters, and yet the tracing diagram is made on a sheet of notebook size paper.

The authors wish to acknowledge the assistance of Mr. Leslie L. Foldy in checking the expressions presented.

Analyses of Mixtures of Light Gases by Infra-Red Absorption

NORMAN D. COGGESHALL AND ELEANOR L. SAIER
Gulf Research and Development Company, Pittsburgh, Pennsylvania
(Received January 17, 1946)

A discussion is given of the application of infra-red absorption methods of analysis for light gases which do not obey Beer's law of absorption due to pressure broadening. Some of these gases do not obey Beer's law in the pure state or when contaminated by foreign gases. The method of analysis depends upon the nature and intensity of the pressure broadening effect of the different components in the sample upon each other. Data are presented showing the nature of some of these effects and illustrating the accuracy obtainable for certain types of analyses. The instrumentation used in routine gas analyses by infra-red in this laboratory is described.

IN the last few years there has been an intensified application of infra-red absorption spectrometry to problems of chemical analysis.[1–4] The technique is now being successfully applied to the analysis of mixtures in the gas and liquid phase and to identification and studies of molecular structure for materials in all phases. One application that has been of particular economic importance recently is gas analysis. Brattain, Rasmussen, and Cravath[2] published what was perhaps the first detailed procedure for multicomponent gas mixtures wherein they were interested in the four-carbon hydrocarbons. The same technique, which is based on the principle that the compounds present obey Beer's law of absorption, may be applied to other gas mixtures. These may be of C_5 hydrocarbons, i.e., pentanes and pentenes as well as many others.

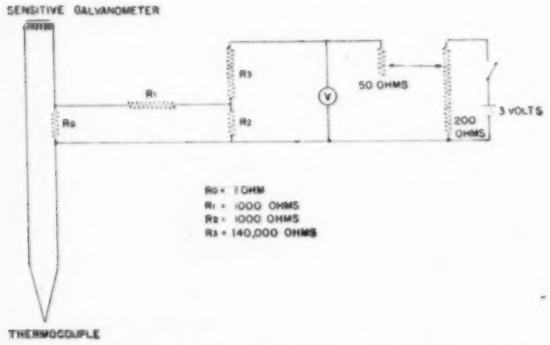


FIG. 1. Null circuit used to measure radiation intensities.

[1] Barnes, Gore, Liddel, and Williams, *Infra-Red Spectroscopy* (Reinhold Publishing Corporation, New York, 1944).
[2] Brattain, Rasmussen, and Cravath, J. App. Phys. **14**, 418 (1943).
[3] Norman Wright, Ind. Eng. Chem. Anal. Ed. **13**, 1 (1941).
[4] April–May, 1944 issue of *Transactions of the Faraday Society* devoted to "The Application of Infra-Red Spectra to Chemical Problems."

Beer's law of absorption states that

$$I = I_0 e^{-\alpha l p}, \tag{1}$$

where I is the intensity of light of a particular wave-length transmitted by a sample of material, I_0 is the intensity of light of the same wave-length incident upon the sample, α is a constant which depends upon the material in the sample and the wave-length under investigation, l is the thickness of the sample, and p represents the concentration of the material in the sample which in case of gas would be the pressure. The constant α may assume different values depending upon what system of units is used for l and p.

For the successful application of the method of Brattain and others[2] to multicomponent mixtures, such that there is appreciable overlapping of the absorption bands of the different components, it is very important that each obey Beer's law of absorption. This is true for the mixtures named above as well as for others in both the gas and liquid phase. However, for some of the light gases such as CO, CO_2, SO_2, CH_4, N_2O, HCl, etc., Beer's law is not obeyed and a somewhat different technique must be used. For these gases, the light absorption depends upon the total pressure in the absorption cell as well as the partial pressures of the absorbing materials. For example, a cell, containing 10 cm of Hg pressure of methane alone, will not absorb the same as if the cell had in addition to the same pressure of methane a partial pressure of some gas such as H_2 which does not absorb in the infra-red. More will be said about this later.

This anomalous pressure dependence of the absorption of the light gases has been known for

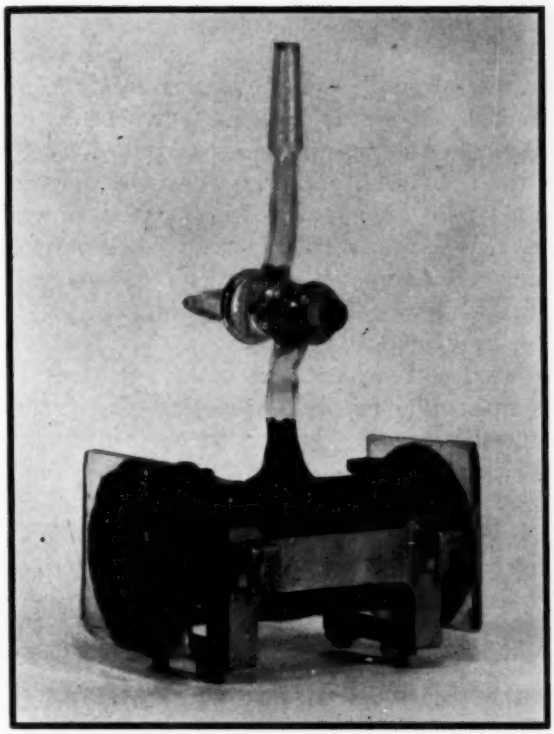


FIG. 2. Absorption cell assembly.

some time and various investigators have studied the effect in different materials.[5–9] D. M. Dennison has made a theoretical study of the factors affecting the shape and intensities of infra-red absorption lines.[10] His results are based on the assumption that the principal factor in the broadening of an absorption line is the limitation of the length of wave train a molecule may absorb due to its perturbation by thermal collisions. He found that the absorption depends upon the number of molecules per unit volume, the temperature of the gas, and the effective molecular diameters. Experimental work was done by Cross and Daniels on the absorption of nitrous oxide as influenced by foreign gases. In line with Dennison's results they analyzed their data to obtain "optical" collision diameters for the foreign gases.

[5] Eva. v. Bahr, Ann. d. Physik **29**, 780 (1909); Physik. Zeits. **33**, 585 (1910); Verh. d. D. Phys. Ges. **15**, 710 (1913).
[6] G. Hertz, Verh. d. D. Phys. Ges. **15**, 673 (1913).
[7] C. Schafer and F. Matossi, *Das Ultra-Rote Spectrum* (Verlagsbuchhandlung Julius Springer, Berlin, 1930), §30.
[8] P. C. Cross and F. Daniels, J. Chem. Phys. **2**, 6 (1934).
[9] Other references can be found in article: Nielsen, Thornton, and Dale, Rev. Mod. Phys. **16**, 307 (1944).
[10] D. M. Dennison, Phys. Rev. **31**, 503 (1928).

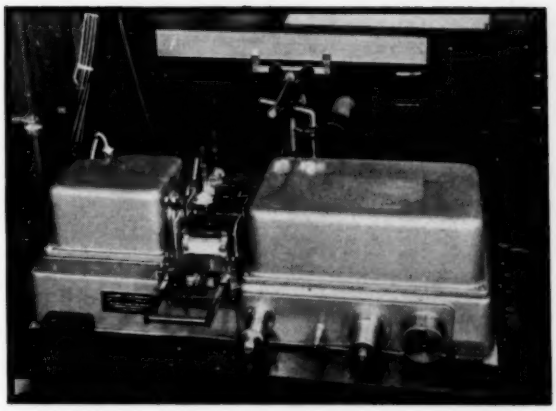


FIG. 3. Infra-red spectrometer used with added cell positioning equipment.

The data given below will illustrate the pressure broadening and how it is encountered in the making of actual analyses for some of the light gases. Several cases, for example, in the petroleum industry wherein it is desirable to make such analyses, are: hydrogen rich gases such as are encountered in hydroforming, flue gas from a catalytic cracking unit, and burn-off gases from catalyst regeneration. Others working in the application of infra-red spectroscopy have doubtlessly encountered the same problems and perhaps evolved essentially the same solution. In that case the material of this article will illustrate the method and apparatus as used by us, the particular procedure, the accuracy obtainable by this procedure, and data of general interest with regards to pressure broadening.

INSTRUMENTAL DETAILS

The infra-red spectrometer used in this work was a Perkin-Elmer Model 12A. The thermo-

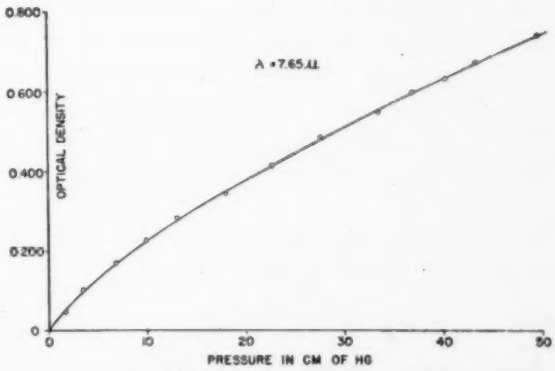


FIG. 4. Optical density *vs.* pressure for methane.

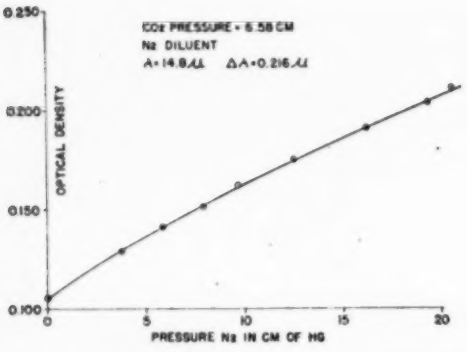


FIG. 5. Influence of N_2 on absorption of constant partial pressure of CO_2.

couple signals were fed into a Leeds and Northup, HS No. 2284–6, 0.05 μv/mm sensitivity galvanometer. A null method was used rather than a direct photometric method wherein the galvanometer deflections are read. For this a controlled voltage was fed into the circuit containing the thermocouple and galvanometer in a direction opposite to the voltage signals due to radiation falling on the thermocouple. The controlled voltage was not measured directly but by measuring a large voltage to which it was in constant proportion. The circuit for the null method is to be seen in Fig. 1. Here the bucking voltage originates from the 3-volt battery which is connected directly across a potentiometer which is of 200-ohm resistance. The moving contact on the 200-ohm potentiometer provides a fairly good control but an additional fine control is available through the moving contact on the 50-ohm potentiometer. The resulting voltage tapped off by this variable resistance system is read on the voltmeter which is a Weston, Model 45, and which has 3 volts full scale as one of its ranges.

It is to be noticed that the circuit integral with and to the left of the voltmeter in Fig. 1 is continuous in the sense of no moving or variable contacts. Therefore, the voltage across R_0 is always a constant fraction of that across the voltmeter. If V represents the voltage across the voltmeter and V_0 the impressed voltage across R_0, we may deduce for the case of balance by simple applications of Ohm's law

$$V_0 = \tfrac{1}{2}(R_0/R_3)\,V, \qquad (2)$$

when R_1 and R_2 are equal and $R_3 \gg R_1$. With the resistances as indicated on Fig. 1 it is possible to introduce a bucking voltage across R_0 up to 3.6 microvolts. In operation, when a signal to the thermocouple is to be determined, the mechanical zero for the HS galvanometer is first observed on a ground glass scale with a shutter in the light path and with the potentiometers adjusted to give zero voltage across the voltmeter. When the shutter is removed the potentiometers are adjusted to bring the galvanometer deflection back to the starting point. The voltage appearing on the voltmeter is then read and recorded as representative of the signal to the thermocouple. The system is very reproducible, successive readings giving values which have variations not larger than 0.2 percent. The 3-volt battery is in reality a combination of two Burgess 4F2H batteries connected in parallel for longer life. The batteries, resistances, and potentiometers are all mounted in an iron-walled case. This is important to protect the circuit from magnetic disturbances, which in a network of these impedances can cause serious disturbances.

The infra-red absorption cells used in this work have glass walls, an inside length of 9.5 cm, and have rocksalt windows. They are equipped with a stopcock, an inner part of a ground glass joint, and an exterior aluminum mounting frame. The bodies of the cells were painted black to minimize the amount of radiation of visible wave-lengths which may enter the monochromator chamber. This assembly is shown in Fig. 2. The aluminum frame has adjusting screws which allow the cell to be positioned as desired relative to it. Also the frame is equipped with slotted positioning brackets so that it may be clamped securely in place in the spectrometer. Two absorption cells

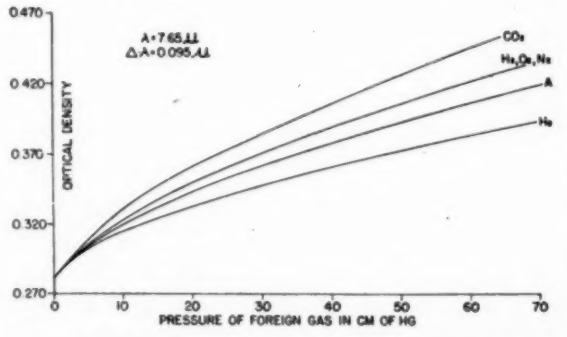


FIG. 6. Influence of gases: He, A, H_2, O_2, N_2, and CO_2 on absorption of constant partial pressure of CH_4.

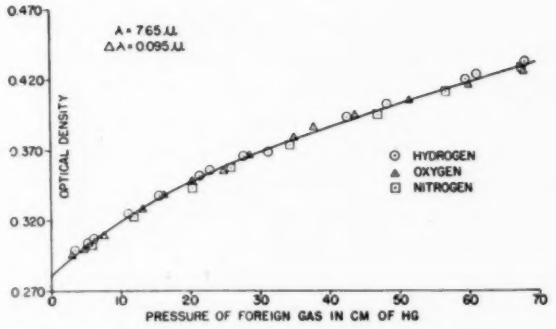


FIG. 7. Detail of influence of H_2, O_2, and N_2 on absorption of methane.

are used, one for the sample and one for comparison or incident energy readings. The two cells are mounted on a carriage mechanism which moves in such a manner that either cell may be moved into the optical path.

The instrument with the cell moving carriage with cells in place can be seen in Fig. 3. The carriage is equipped with a bar which the operator can use to position the cells. Adjustable stops insure that the cells always arrive in the same position in the light path. To remove the cells all that is necessary is to loosen two thumbscrews. The cells are evacuated and filled to measured pressures on a gas handling system separate from the instrument. Connected permanently to the thermocouple case in the instrument is a small mercury diffusion pump which is operated periodically to maintain a good vacuum in the presence of the radiation receiver. The apparatus here has been in use for some time for the analyses to be described as well as in the routine analysis of mixtures of the seven common C_4 hydrocarbons. For this, as well as for the present application, it has been very satisfactory and allows considerable speed of data taking. When it is desired to change the pressure in the absorption cell it is merely removed and engaged by means of its standard taper ground glass joint, into a gas handling system designed specifically for handling of routine samples.

In making blends and calibration mixtures for the analyses to be described as well as for other work of this type, it is important that the different components be thoroughly mixed. If the total pressure of the mixture in the absorption cell did not exceed 25 cm of Hg it was mixed in a separate, constant volume chamber. This chamber was equipped with a thin steel diaphragm which is moved back and forth to mix the gases by means of exterior magnets. When the total pressure of the mixture in the cell exceeded 25 cm of Hg the mixing was done in the absorption cell itself. This was possible as the cell was equipped with a thin steel vane bent to fit the inside curvature of the cell. When the cell was in the operating position the vane lay flat against the bottom, out of the optical path. Mechanical motion of the cell such as shaking or rotating caused the vane to move back and forth, thoroughly mixing the gases.

DATA, CALIBRATION, AND PROCEDURE

For a gas obeying Beer's law a straight line curve is obtained when the observed optical density D ($D = \log I_0/I$) is plotted as a function of the pressure of the gas in the absorption cell. As stated above, some of the light gases do not obey Beer's law and data illustrating this may be seen in some of the following figures. In Fig. 4 is seen a graph of the optical density plotted against pressure for methane alone. As may be noticed, it is far from a straight line. It will be observed shortly that the general appearance of this plot is similar to those obtained when a constant pressure of methane is used and the partial pressure of a non-absorbing foreign gas is varied. In Fig. 5 may be seen a curve for the optical density for CO_2 plotted against partial pressure of added N_2. It is observable that the effect of the N_2 is to increase the absorption very greatly. All the non-absorbing gases used were tested for absorption at the wave-lengths used. In each case they showed zero absorption.

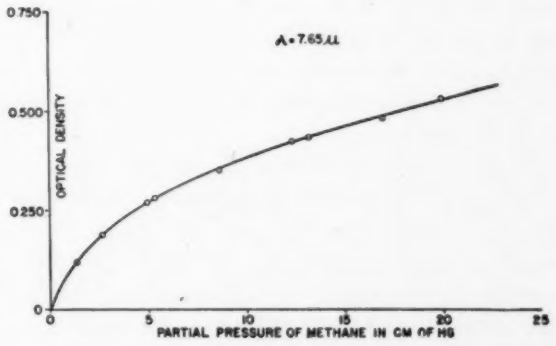


FIG. 8. Calibration curve constructed for analysis for methane.

Figure 6 shows the effect of various foreign gases on the absorption due to methane. The graphs shown in this figure were obtained by drawing smooth curves through the data obtained for each type of foreign gas. The amount of scattering of the data about the curve in each case is about the same as will be seen in Fig. 7. Here the ordinate represents the observed optical density and the abscissa represents the partial pressure of the foreign gas added. The data were all taken at the wave-length $\lambda = 7.65\ \mu$ where methane has its strongest absorption. The slits were used at a setting which gave a spectral slit width of $0.095\ \mu$ at this point. The partial pressure of methane was always constant, being 13.15 cm of Hg. The foreign gases added, namely, helium, argon, oxygen, nitrogen, hydrogen, and carbon dioxide, have no infra-red absorption at this wave-length and the large change in optical density produced is a manifestation of the pressure broadening effect mentioned earlier. It is to be noticed that the effects of the various gases are different, some of them being responsible for a greater increase in optical density than others. A study of these data in terms of the pressure broadening theory of Dennison and others has been made.[11] In Fig. 6 the effect due to nitrogen, hydrogen, or oxygen is shown as one curve; Fig. 7 shows in more detail the effects of these latter gases.

For the analyses in which we are interested here Fig. 7 is of primary importance. This shows that the effects of hydrogen, nitrogen, and oxygen as added gases are nearly the same. Thus it is possible to construct a single empirical calibration curve which will serve for the analysis of methane in the presence of any one or mixture of the three. This is important as the off gases for which the method is used vary with regards to the relative concentrations of H_2, N_2, and O_2.

Figure 8 shows the calibration curve constructed for this analysis. For every determination the total pressure in the absorption cell was the same, i.e., 80 cm of Hg. The points were obtained by blending known amounts of methane and hydrogen together, admitting to the cell, adjusting the pressure to 80 cm, and determining the optical density. Figure 8, therefore, represents the measured optical density *versus* partial pressure of methane, the total cell pressure being constant.

With this curve the actual analysis of sample consists only of the steps of: admitting the sample into the cell to the pressure of 80 cm, determining the optical density, and from the curve reading off the partial pressure. Table I gives some results for a series of synthetic samples analyzed by means of the previously constructed calibration curve. The table shows the percent of methane as blended, the percent as determined by the analysis, the nature of the remainder of the sample, and the percent difference in terms of total sample between the synthetic and observed concentrations. The table not only illustrates the obtainable accuracy but also shows it to be independent of the nature of the remainder of the sample. Hundreds of samples have been analyzed for methane by this method and the results at all times have been satisfactory. The time required for a complete analysis, including introducing the sample, taking the data, and making the calculations, is about 15 minutes under routine conditions.

An analysis somewhat similar to the one just

TABLE I. Results for synthetic samples for methane analysis.

Sample No.	% CH_4 (synthetic)	% CH_4 (observed)	Δ%	Nature of remainder of sample
1	14.5	14.7	0.2	48% H_2; 52% N_2
2	7.6	7.7	0.1	N_2
3	6.00	6.05	0.05	33% H_2; 67% O_2
4	22.5	22.3	0.2	O_2
5	4.90	4.95	0.05	Air
6	3.71	3.76	0.05	H_2
7	1.29	1.36	0.07	Air
8	0.83	0.85	0.02	3.9% Air; 96.1% H_2

TABLE II. Composition of blends used to construct calibration curves for flue gas analyses.

Blend No.	% CO_2	% SO_2	% CO	% Air
1	9.5	2.0	3.4	85.1
2	12.7	2.0	6.6	78.7
3	3.9	1.7	8.7	85.7
4	2.5	1.1	10.8	85.6
5	1.6	0.7	10.9	86.8
6	7.6	1.3	4.7	86.4
7	7.0	0.4	1.4	91.2
8	4.9	1.2	2.9	91.0
9	6.5	0.6	1.3	91.6

[11] N. D. Coggeshall and E. L. Saier, Bull. Am. Phys. Soc. **21**, 17 (1946).

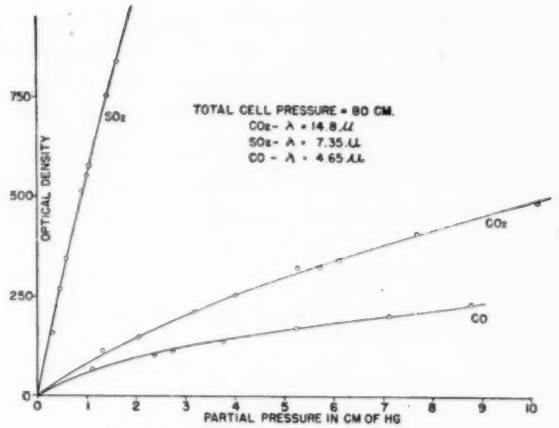


FIG. 9. Calibration curves for analysis of flue gas for CO_2, SO_2, and CO.

described, but involving more components, is the analysis of flue gases and burn-out gases from catalyst regeneration for CO_2, CO, and SO_2. In this case also the remainder of the samples varies, containing different amounts of O_2, H_2, and N_2. Since for these samples the major part is generally a mixture of these three and since the concentration of any one of the three gases, CO, CO_2, or SO_2, is generally less than 10 percent it was decided to try the same method as used for the methane analysis.

A calibration curve for each of the three compounds CO, CO_2, and SO_2 was constructed as was done for methane. In constructing these, blends of the three gases and of air were made and examined as before. The wave-length used for CO_2 was 14.8 μ, for SO_2 it was 7.35 μ, and for CO it was 4.65 μ. Each gas has a maximum of absorption at about these wave-lengths and there is no overlapping between their absorption bands at these points. As before the blend was admitted to the absorption cell to the same pressure each time, this being 80 cm of Hg. In making the blends, care was taken to see that the concentrations of CO, CO_2, and SO_2 were randomly distributed, that is to say: the concentration of any one of the three in any one of the blends was independent of the concentration of the other two. Table II gives the concentrations of the blends used for constructing the calibration curves.

The calibration curves are shown in Fig. 9. Here the uniformity and smoothness of the curves attests to the independence of the absorption of the three gases. In this figure the optical density observed at each wave-length was plotted against the partial pressure of the gas which absorbed at that point. The slits were set to a spectral slit width of 0.216 μ for readings of 14.8 μ; 0.117 μ for readings of 7.35 μ; and 0.085 μ for readings of 4.65 μ. Smaller slit widths could be used but the ones employed allowed the Globar element to be operated with about 150-watt supplied power. At this power level the life of each element is quite long. The shape of the curves indicates the deviations from Beer's law to be the same type as for methane.

Although the smoothness of the calibration curves is a manifestation of the internal consistency of the data, the usefulness and accuracy of the curves were tested by a further set of synthetic blends. The results are shown in Table III. Here it may be seen that the average error in terms of total sample is about 0.1 percent. The accuracy for CO is somewhat surprising in view of the flatness of the calibration curve for this compound.

As flue gas analyses may be used for process

TABLE III. Comparison between blended synthetic samples and concentrations measured by infra-red absorption for flue gas analyses.

Sample No.	Compound	Synthetic	Observed	% Dev.
1	CO_2	10.5	10.4	0.1
	SO_2	2.1	2.3	0.2
	CO	4.2	3.9	0.3
2	CO_2	6.7	6.9	0.2
	SO_2	1.4	1.4	0.0
	CO	9.4	9.5	0.1
3	CO_2	6.2	6.2	0.0
	SO_2	1.9	1.9	0.0
	CO	4.0	4.1	0.1
4	CO_2	4.0	4.2	0.2
	SO_2	1.3	1.3	0.0
	CO	2.6	2.5	0.1
5	CO_2	2.7	2.8	0.1
	SO_2	1.0	1.0	0.0
	CO	5.2	4.8	0.4
6	CO_2	1.7	1.9	0.2
	SO_2	0.6	0.6	0.0
	CO	3.4	3.1	0.3
7	CO_2	7.0	7.1	0.1
	SO_2	1.9	1.9	0.0
	CO	6.1	6.3	0.2
8	CO_2	4.4	4.6	0.2
	SO_2	1.2	1.2	0.0
	CO	4.0	3.9	0.1

control it is important that the complete procedure be as simple as possible. This is important since it allows them to be done on night shifts by relatively untrained personnel. The equipment and method as described above have been used for process control analysis for a fluid catalytic cracking unit for some time and hundreds of samples have been handled for this purpose. One operator is sufficient to carry out all steps and the complete analysis, including calculations, can be done on a routine basis in about 35 minutes or less. This includes duplicate determinations of the optical densities at each wave-length value.

The authors are indebted to Mr. N. F. Kerr for aid in developing some of the equipment used, to Dr. Morris Muskat for various suggestions in the application of infra-red absorption methods, and to Dr. Paul D. Foote, Executive Vice-President of Gulf Research and Development Company for permission to publish this material.

Methods of Minimizing Lead Loss in Emissivity and Resistivity Determinations

L. MALTER, *RCA Victor Division, Radio Corporation of America, Lancaster, Pennsylvania*

AND

D. B. LANGMUIR,* *Office of Scientific Research and Development, Washington, D. C.*

(Received January 12, 1946)

Two means are described for minimizing the effects of lead loss in determining the electrical properties of thin filaments as a function of temperature. In both cases the filaments are supported at the mid-points of filamentary U's. In one case the U's are made of finer wire than the main filament so that lead loss is reduced. In the other case, alternating current is passed through the U's so that the temperature of the main filament is maintained uniform down to its points of support.

DURING the course of work some time ago on the properties of tantalum filaments,[1] it was desired to extend the measurements to as low temperatures as possible. At these low temperatures difficulties were encountered because of the fact that the cooling effect of the filament supports extended up to and beyond the potential leads. Since in resistivity determinations it is essential that the filament temperature be uniform over the region (between the potential leads) whose electrical constants are being measured, it becomes necessary when following standard technique to use inordinately long filaments, unless they are of other than gossamer-like thickness.

To obviate this difficulty, recourse was had to one of the two expedients which are described below. In the first of these (see Fig. 1) the filament, whose properties are to be determined, is welded at each end to the mid-point of an inverted U of considerably smaller size wire. By proper proportioning of the diameters of filament and U's, it is possible to secure uniform temperature of the filament not only over the region between the potential leads but in fact almost to the filament supports. A little experimental work may be necessary to arrive at the correct ratio of diameter of filament wire to U wire for securing satisfactory operation over the range of interest. In our work, a ratio of two to one was found satisfactory.

The second method illustrated in Fig. 2 is preferable in that no preliminary experimental work is required, and in that the filament temperature can be made uniform to its very ends for all temperatures. The inverted U's which support the filament in this case are made of the same size wire as the filament itself. With only the d.c. from the battery flowing through the U's to the filament, the U's run considerably cooler

* Formerly with RCA Laboratories, Princeton, New Jersey.

[1] L. Malter and D. B. Langmuir, Phys. Rev. **55**, 743 (1939).

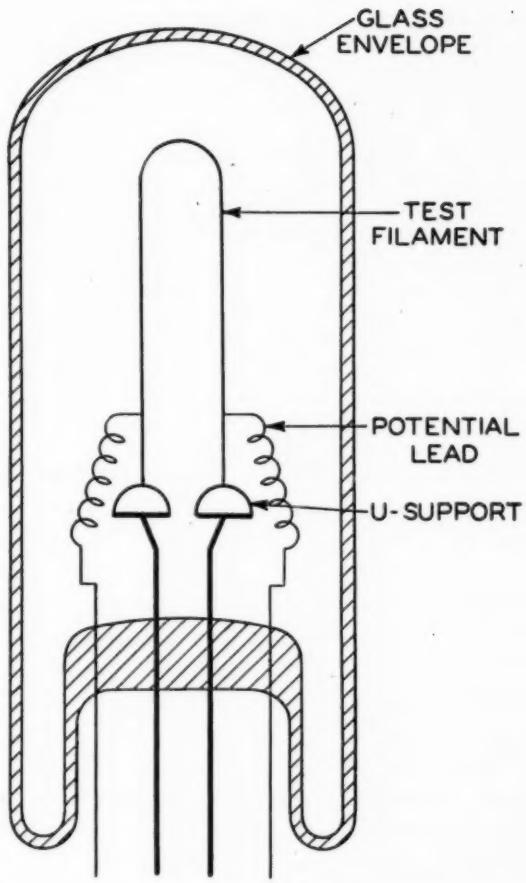


FIG. 1. Fine wire U supports for minimizing lead loss effects.

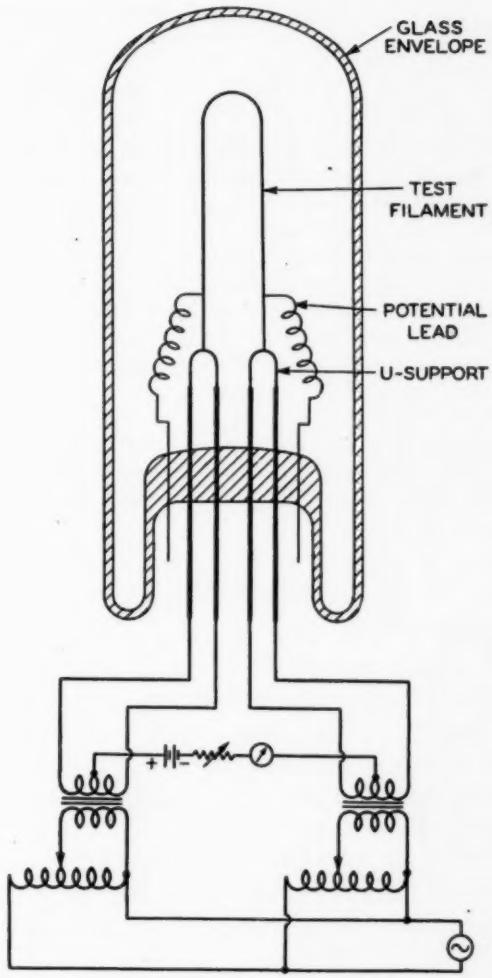


FIG. 2. Auxiliary heat method for minimizing lead loss effects.

than the filament, thus resulting in a cooling of the latter to some distance from their points of support. However, by passing alternating current through the U's by means of external transformers, the temperature of the U's may be raised so that the main filament temperature is uniform right up to its points of support, the alternating current through the U's being controlled by means of small autotransformers as shown. Since no alternating current flows through the filament itself, the determination of the electrical characteristics by d.c. methods between the potential leads is unaffected.

By use of these methods, it has been possible to extend the range of measurements to the limit of visibility. It appears that by a process of extrapolation, it should be possible to secure data beyond this limit.

Theory and Application of the Parallel Plate Plastometer*

G. J. Dienes and H. F. Klemm
Bakelite Corporation, New York, New York

A method has been established for the measurement of the viscosity of high polymers at low rates of shear in the range 10^4 to 10^9 poises using a parallel plate plastometer. This is based on a mathematical criterion for separating the viscous portion of the deformation from the "elastic" and "delayed elastic" components. Experimentally, the plate separation is measured at a given temperature as a function of time. The theory furnishes a relation, which is also the criterion for predominantly viscous deformation, between viscosity, plate separation, applied load, and time. This relation, a modified form of Stefan's equation, is used for calculating the viscosity from the experimentally observed quantities. The method has been applied to polyethylene and vinyl chloride-acetate resin compounds. The viscosity-temperature behavior of these materials is shown to be simple over the temperature range studied; that is, log viscosity varies linearly with the reciprocal of the absolute temperature. Data are presented which show that polyethylene resins and polyethylene resin-paraffin wax mixtures follow Flory's relation; that is, log viscosity varies linearly with the square root of the weight average molecular weight. Accordingly, the parallel plate plastometer offers promising possibilities for the empirical determination of the weight average molecular weight of these materials. Data are also presented on plasticized vinyl chloride-acetate resin systems which point to a close parallel between the effects of increasing temperature and increasing plasticizer concentration.

INTRODUCTION

ONE of the most important properties of plastic materials is their ability to flow under the action of applied stresses at elevated temperatures. The quantitative evaluation of flow behavior in terms of applied stress and temperature is of great practical importance and theoretical interest.

It is well known that the behavior of plastic materials under the action of stresses is very complex because elastic and viscous deformations take place simultaneously. The nature of stress-strain and deformation-time (constant load) curves has received the attention of many investigators in recent years.[1] It is generally agreed that a combination of elastic and viscous properties gives rise to the characteristic deformation-time or "creep" curves of high polymers.

A typical deformation-time curve is illustrated in Fig. 1. A quantitative separation of this curve into its elastic and viscous components has not yet been accomplished, although a

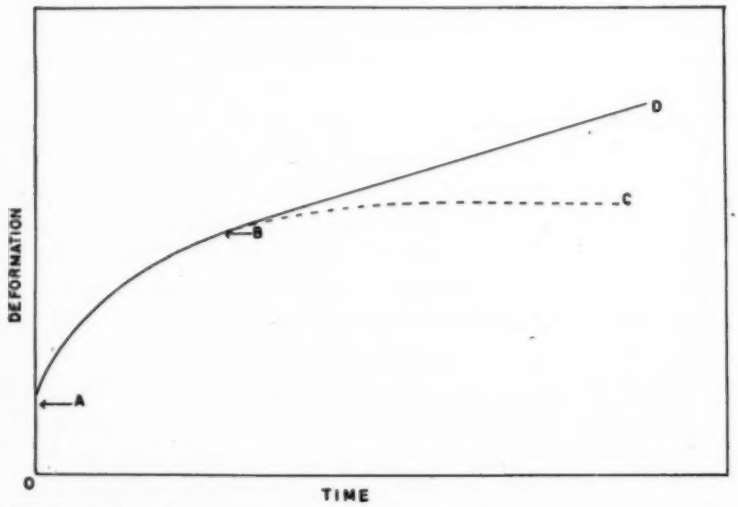


FIG. 1. Typical deformation-time curve.

* Presented at the Annual Meeting of the Society of Rheology, October 26 and 27, 1945.

[1] M. Reiner, *Ten Lectures on Theoretical Rheology* (Nordemann Publishing Company, New York, 1943). R. F. Tucket, Chem. Ind. 430 (Nov. 1943). D. R. Morey, Ind. Eng. Chem. **37**, 255 (1945). G. W. Scott Blair, *Survey of General and Applied Rheology* (Pitman Publishing Corporation, New York, 1944). T. Alfrey, Quart. App. Math. **3**, 143 (1945); J. Chem. Phys. **12**, 374 (1944).

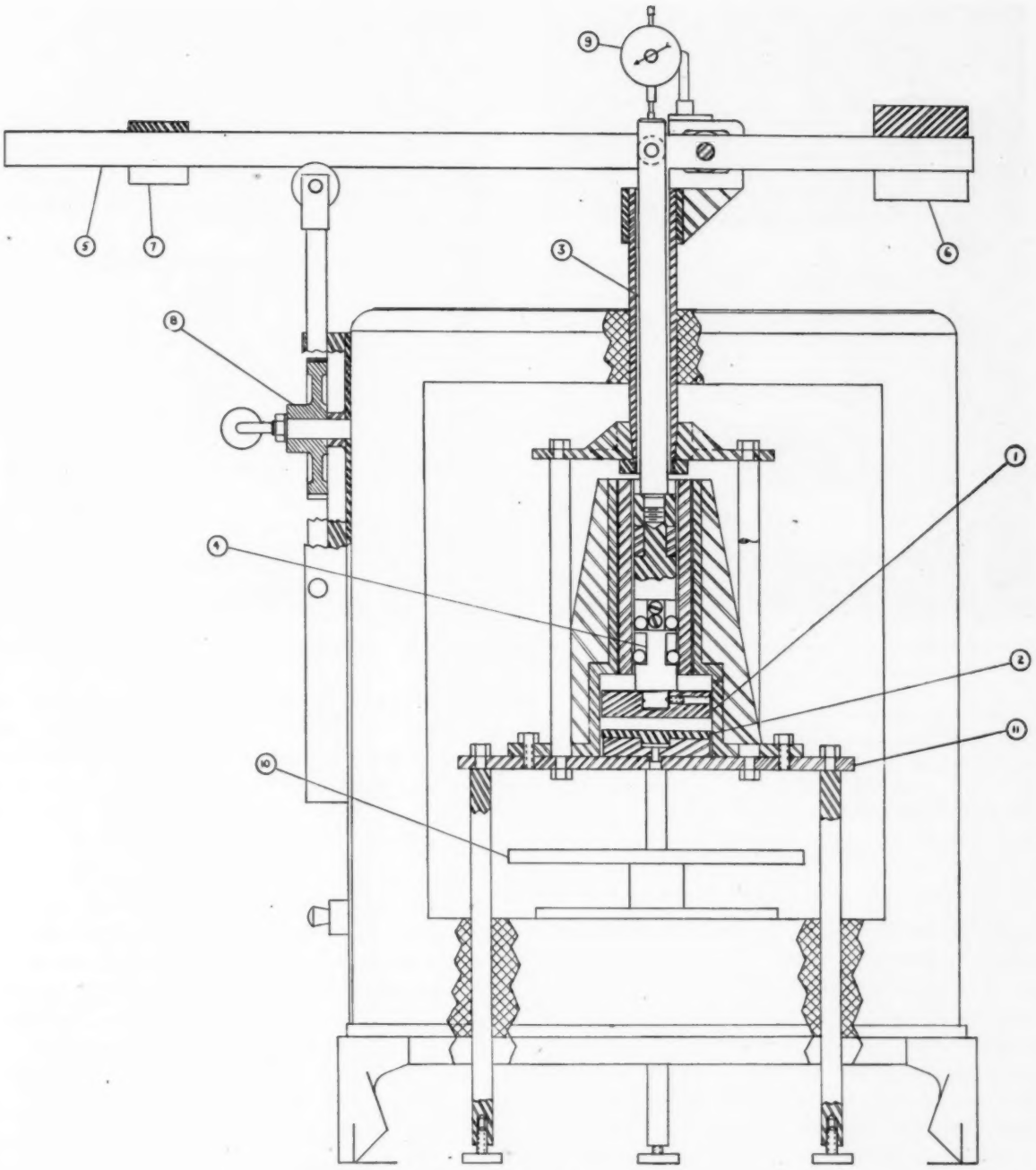


FIG. 2. Parallel plate plastometer.

qualitative separation into regions of predominance is possible. With reference to Fig. 1, this is done as follows:

(1) The first region is represented approximately by the portion of the curve OA. This represents a predominantly "elastic" deformation which takes place rapidly and is recovered similarly when the load is released.

(2) The second region is represented approximately by the portion of the curve AB, and is often referred to as the "delayed elastic" deformation. This deformation is characterized by a rapidly changing slope and may take place over a considerable period of time. It approaches a constant value as indicated by the broken line BC. The part BC can be observed only when no viscous deformation takes place.

(3) Region three is represented by a straight line, BD.

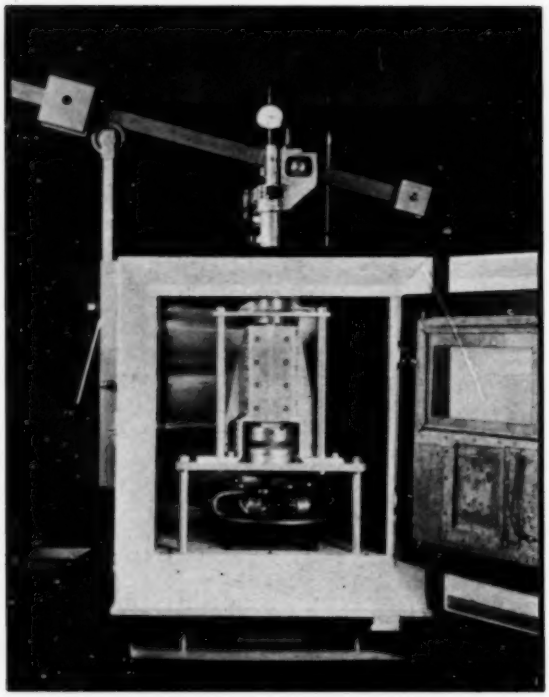


FIG. 3a. Plastometer showing 4-in. diameter plates separated.

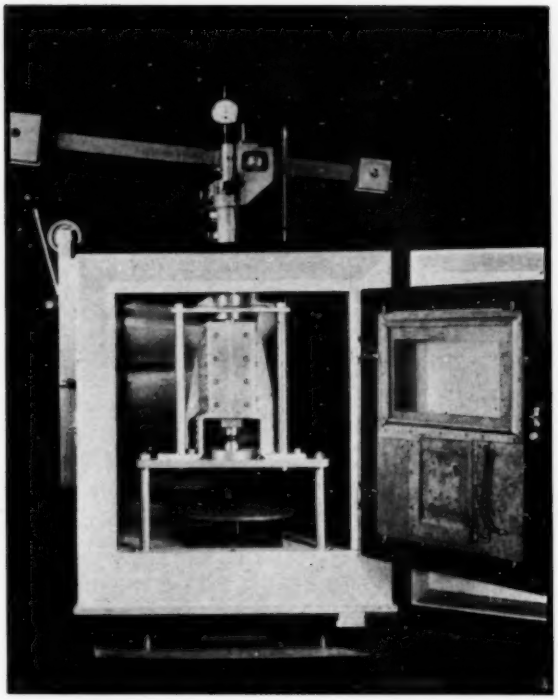


FIG. 3b. Deformation under load test (A.S.T.M. D 621-B).

It is generally agreed that this portion of the curve represents predominantly viscous deformation which is non-recoverable, and continues until the load is removed, the specimen is ruptured, or the test conditions are altered. This part of the curve is characterized by a constant slope. This does not apply, of course, at stresses and temperatures where a permanent deformation does not occur.

The three regions described above represent a rough separation of the deformation-time curve into its component parts. For a further understanding of the mechanical properties of plastics, it is essential to carry out this separation quantitatively. The final aim is to describe the complex mechanical behavior of high polymeric materials in terms of such fundamental constants as viscosity, modulus of elasticity, relaxation times, etc.

This paper describes an attempt to apply the parallel plate plastometer to the measurement of the predominantly viscous deformation (region three) free from "elastic" and "delayed elastic" effects.

MECHANICAL DETAILS

The improved parallel plate plastometer used for the evaluation of viscosity of plastics is shown in front elevation (sectioned for clarity) in Fig. 2. The instrument consists of two circular parallel plates, both either 4.00 or 1.125 in. in diameter, the upper (see 1, Fig. 2) movable, the lower (2) stationary. The upper plate is attached to the vertical shaft (3) which is aligned and guided in the low friction ball-bearing assembly illustrated (4). The test load is applied by means of the beam (5), mounted at the upper end of the shaft (3), and the two movable weights (6 and 7). The dial indicator (9), which makes contact with the loading shaft, is fastened to the frame or reference system of the plastometer, and indicates the plate separation at any instant. A hand-operated rack and pinion mechanism (8) on the left-side of the oven permits the application of the test force without opening the oven door. The plastometer is supported by three legs, fitted with leveling screws, which project through the bottom of the oven.

Figures 3a and 3b are photographs of the plastometer which further illustrate its mechanical details. Figure 3a shows the apparatus with both the inner and outer oven doors open. The turntable (10, Fig. 2) underneath the plastometer proper, supports several clamps which hold the specimen assemblies together when preheating is done outside of the test area of the plates.

Figure 3b shows the equipment fitted with special plates (1.125 in. in diameter, one square inch in area) which are required in some deformation under load tests (A.S.T.M. D 621, Method B). These plates are also used when the plastometer is operated as Type 1 (see Theory), i.e., the plates have the same area as the ends of the specimen.

The important advantages of the parallel plate plastometer described above are listed below:

(1) Low friction losses. All movable points of contact are provided with ball bearings to minimize friction.

(2) Wide range of test loads (up to about 60 kilograms (132 lb.)).

(3) External application of test load.

(4) Dial indicator is attached to the plastometer frame and is located outside of the oven for accurate trouble-free operation at all test temperatures.

(5) Small specimen door, located in inner of two oven doors, minimizes temperature changes due to sample insertion and removal.

(6) Parallel plates removable for ease of cleaning and permitting change from one type of plates to another.

A more recent model of this instrument is equipped with a loading beam calibrated in units of 0.1 k.g.s.

EXPERIMENTAL PROCEDURE

Calibration of the Plastometer

The calibration of the plastometer was accomplished at room temperature by measuring the force transmitted by the upper plate to a platform balance. The applied force was varied by changing the position of the weights on the loading beam. The balance was placed in the oven beneath the main horizontal support member (11, Fig. 2) of the plastometer. This force was transmitted to the balance by a connecting rod which passed through the hole in the main horizontal member. Two threaded disks were attached to either end of the rod, and the assembly adjusted for length to obtain proper equilibrium position of the balance.

The calibration data in graphical form are shown in Fig. 4. It will be noted that both ranges are characterized by linear relations between load and the distance D_L. Furthermore, the observed calibration agreed very closely with that obtained by calculation.

Test Specimens

The test specimens were right circular cylinders, consisting of a stack of disks 0.63 in. (1.60 cm) in diameter, cut from sheet stock 0.040 to 0.080 in. (0.101 cm to 0.202 cm) thick. The specimen mass in grams was equal to the mass of an integral number of disks closest to twice the density of the material. Such a specimen has a height of about 0.394 in. (1.0 cm) and a volume of about two cm^3.

Test Procedure and Calculations

An outline of the test procedure for determination of viscosity by the parallel plate plastometer is given below.

(1) Adjust the oven to the desired test temperature, and allow sufficient time (at least one hour) to insure thermal equilibrium in the apparatus.

(2) Position the loading weights along the lever arms to give the desired test load.

(3) Set the dial gauge so that it indicates minus two divisions (one div. = 0.001 in.) when the plates are together under test load (no specimen). This corrects for the thickness of the two one-mil sheets of cellophane which are

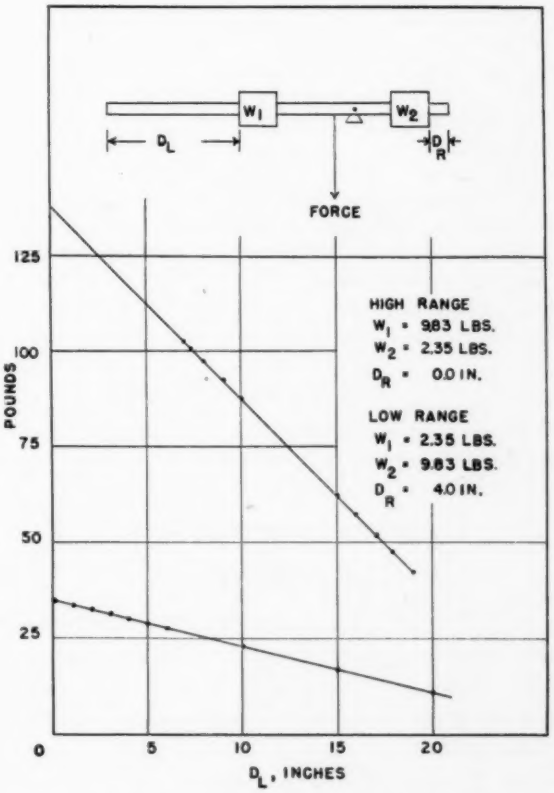


FIG. 4. Plastometer load system calibration.

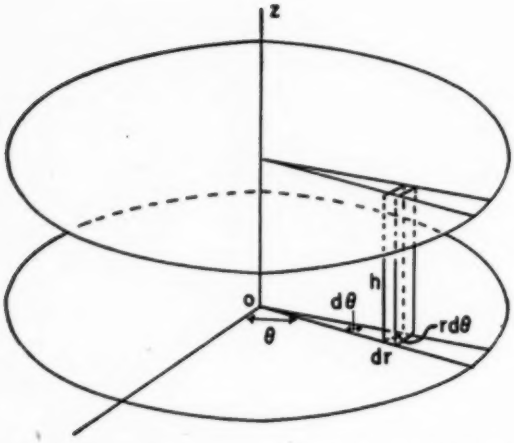


FIG. 5. Schematic diagram of plastometer illustrating coordinate system.

placed on the two ends of the specimen, and fixes the zero point for specimen height readings.

(4) Center the specimen (weighed to the nearest milligram) between two 4×4×0.001 in. pieces of plain cellophane, and insert this assembly centrally between the parallel plates. If preheating of the specimen is required, this may be done either in the plastometer or in special clamps provided for this purpose.

(5) Apply the test load and read the separation, h, of the two parallel plates on the dial gauge to an accuracy of ±0.0002 (0.2 of a division) as a function of time.

(6) Plot on linear coordinate paper, as the test proceeds, the reciprocal of the fourth power of the plate separation, $1/h^4$, expressed in cm^{-4}, as ordinate against the time in seconds. A table of $1/h^4$ values (in cm^{-4}), in the range $h=0.010$ in. to $h=0.150$ in., facilitates this operation.

(7) Continue the test until a straight line of slope, m, is defined in the range of h corresponding to $R:h=10$ (R = radius of specimen) and $h=10$ divisions (0.010 in.). Loads and temperatures must be adjusted to satisfy these requirements. Otherwise, the data are not simply interpretable.

(8) Calculate the viscosity, η, from the following equation:

$$\eta=8.21\times10^6(W/mV^2), \tag{16}$$

where η = viscosity, poises, W = applied load, kilograms, m = slope of plot of $1/h^4$ *vs.* time, cm^{-4} $sec.^{-1}$, and V = volume, obtained from specimen mass and density, cm^3.

THEORETICAL CONSIDERATIONS

Preliminaries

In any physical test certain measurements are made, such as length, time, temperature, etc. From these measurements, it is desired to determine quantities which are characteristic of the material. The purpose of theory is to express the physical law which relates the variables measured in the test with the properties of the material.

In the parallel plate plastometer tests, interest is focused on the flow characteristics of the material. Probably the most important constant which describes the flow characteristics of a material is its viscosity.

In the case of the parallel plate plastometer, it is desired to find an equation relating the plate separation, the time, and the viscosity. This problem, which is to find the equation of motion of two plates moving toward each other under the action of a force with a viscous medium between the plates, has been solved,[2] and the solution successfully applied to the parallel plate plastometer.

Derivation of Stefan's Equation from the Fundamental Equation of Viscous Flow

The general equation of motion of a Newtonian fluid of viscosity η (neglecting body forces such as gravity) is given by the vector equation

$$\rho\frac{\partial\mathbf{v}}{\partial t}+\rho\mathbf{v}\cdot\nabla\mathbf{v}=-\nabla p+\eta\nabla^2\mathbf{v}+\tfrac{1}{3}\eta\nabla\nabla\cdot\mathbf{v}, \tag{1}$$

where ρ = density of the fluid, $\mathbf{v}$ = velocity vector, and p = fluid pressure. If the fluid is incompressible, the equation of continuity requires that

$$\nabla\cdot\mathbf{v}=0.$$

If the velocity of the fluid elements is small, the term on the left in the square of $\mathbf{v}$ may be neglected. For slow motion of an incompressible viscous fluid, therefore[3]

$$\rho\partial\bar{\mathbf{v}}/\partial t=-\nabla p+\eta\nabla^2\mathbf{v}. \tag{2}$$

In cylindrical coordinates, Eq. (2) in scalar form becomes

$$\begin{aligned}
-\frac{\partial p}{\partial r}+\eta\nabla^2v_r&=\rho\frac{\partial v_r}{\partial t},\\
-\frac{1}{r}\frac{\partial p}{\partial\theta}+\eta\nabla^2v_\theta&=\rho\frac{\partial v_\theta}{\partial t},\\
-\frac{\partial p}{\partial z}+\eta\nabla^2v_z&=\rho\frac{\partial v_z}{\partial t}.
\end{aligned} \tag{3}$$

Equations (3) are to be solved subject to the

[2] M. J. Stefan, Akad. Wiss. Wien. Math.-Natur. Klasse. Abt. 2., **69**, 713–735 (1874).

[3] L. Page, *Introduction to Theoretical Physics* (D. Van Nostrand Company, New York, 1930), pp. 224–230.

boundary conditions defined by the parallel plate plastometer. In cylindrical coordinates, let the circular parallel plates be the planes (Fig. 5)

$$z=0;$$
$$z=h.$$

The mathematical theory is considerably simplified if the analysis is limited to that part of the data for which the plate separation, h, is small compared to the radius, R, of the specimen cylinder, the ends of which are in contact with the parallel plates.

Because of circular symmetry, v_θ is zero. As long as h is small compared to R, a good approximation is obtained by neglecting the velocity, v_z, normal to the plates, in comparison with the radial velocity, v_r. Thus,

$$v_\theta=0;$$
$$v_z=0.$$

It will be assumed that there is no slippage at the plates, that is

$$v_r=0 \quad \text{at} \quad z=0 \quad \text{and} \quad z=h.$$

As a consequence, the radial velocity, v_r, changes rapidly over the short distance, h, from a finite value in the liquid to zero value on the plates. Under these conditions, then, $\partial v_r/\partial r$ is negligibly small compared to $\partial v_r/\partial z$. From symmetry, $\partial v_r/\partial\theta=0$.

Steady-state flow is assumed, that is

$$\partial v_r/\partial t=0.$$

Thus, the solution will not be applicable at the start of an experiment when transients due to $\partial v_r/\partial t\neq 0$ are not negligible.

With the above simplifications, Eq. (3) becomes

$$\frac{\partial p}{\partial r}=\eta\frac{\partial^2 v_r}{\partial z^2}, \tag{4}$$

where p is now a function of r only.

Experimental work on the parallel plate plastometer indicated that the above approximations are justified as long as $R>10h$. Erratic viscosity measurements, which are difficult to interpret, are obtained when this condition is not obeyed.

Integrating Eq. (4) twice with respect to z and inserting the boundary conditions,

$$v_r=0 \quad \text{at} \quad z=0 \quad \text{and} \quad z=h,$$

gives

$$v_r=\frac{1}{2\eta}\frac{\partial p}{\partial r}(z-h)z. \tag{5}$$

Equation (5) shows that the velocity of the fluid in any plane, $\theta=$constant, varies parabolically in z. It is at this point that M. J. Stefan[2] starts his derivation with the assumption of a parabolic law.

Consider now a cylindrical element (see Fig. 5) extending from one plate to the other, and contained between radii r and $r+dr$ and between two radial planes parallel to z an angle $d\theta$ apart. The flow through a surface element $rd\theta dz$ is $rd\theta v_r dz$. Then U, the flow per unit arc length between the planes $z=0$ and $z=h$ (in the direction of increasing r), is

$$U=\int_0^h v_r dz,$$
$$=\frac{1}{2\eta}\frac{\partial p}{\partial r}\int_0^h (z^2-zh)dz, \tag{6}$$
$$=-\frac{h^3}{12\eta}\frac{\partial p}{\partial r}.$$

Suppose that the plate $z=0$ remains fixed and that the plate $z=h$ remains parallel to it but moves in the z direction with the velocity dh/dt (dh/dt positive in the positive z direction). Because of this motion, the cylindrical element, $drrd\theta h$, will change in volume at the rate

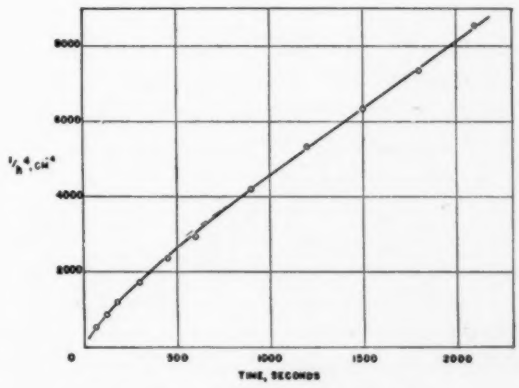


FIG. 6. Typical $1/h^4$ *vs.* time curve for polyethylene at 115°C.

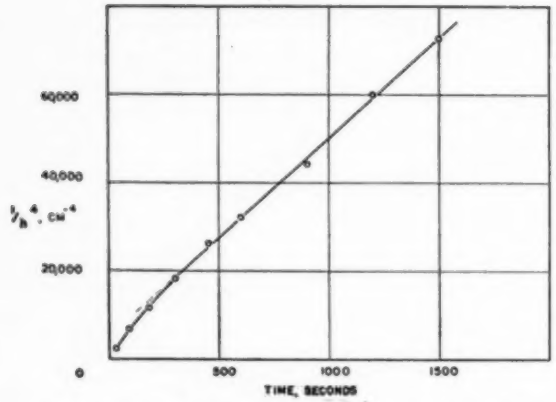


FIG. 7. Typical $1/h^4$ *vs.* time curve for polyethylene at 190°C.

$drrd\theta dh/dt$. Evidently the difference between outflow and inflow must compensate for this change in volume, since the fluid was assumed to be incompressible. If the plate $z=h$ moves toward the plate $z=0$, then the rate of decrease of the volume element must equal the net rate of outward radial flow. Thus,

$$-drrd\theta\frac{dh}{dt}=\frac{\partial}{\partial r}(rd\theta U)dr. \tag{7}$$

Upon substitution of the expression for U and simplification

$$\frac{12\eta}{h^3}\frac{dh}{dt}r=\frac{\partial}{\partial r}\left(r\frac{\partial p}{\partial r}\right). \tag{8}$$

Integration twice with respect to r gives

$$p=\frac{3\eta}{h^3}\frac{dh}{dt}r^2+C'\ln r+C. \tag{9}$$

As p must be finite for $r=0$, C' vanishes. C is evaluated from the condition that at the outer boundary of the specimen, $r=R$, the pressure must be in equilibrium with the constant pressure, Π (usually atmospheric). Thus,

$$C=\Pi-\frac{3\eta}{h^3}\frac{dh}{dt}R^2,$$

and Eq. (9) becomes

$$p=-\frac{3\eta}{h^3}\frac{dh}{dt}(R^2-r^2)+\Pi. \tag{10}$$

The total force acting on the plate in the positive z direction is

$$\int_0^R p2\pi rdr.$$

The total force acting on the plate in the opposite direction is

$$F+\int_0^R \Pi 2\pi rdr,$$

where $F=$applied force (dynes). Therefore,

$$F+\int_0^R \Pi 2\pi rdr=-\int_0^R \frac{3\eta}{h^3}\frac{dh}{dt}\times(R^2-r^2)2\pi rdr+\int_0^R \Pi 2\pi rdr$$

or

$$F=-2\pi\frac{dh}{dt}\frac{3\eta}{h^3}\int_0^R (R^2-r^2)rdr. \tag{11}$$

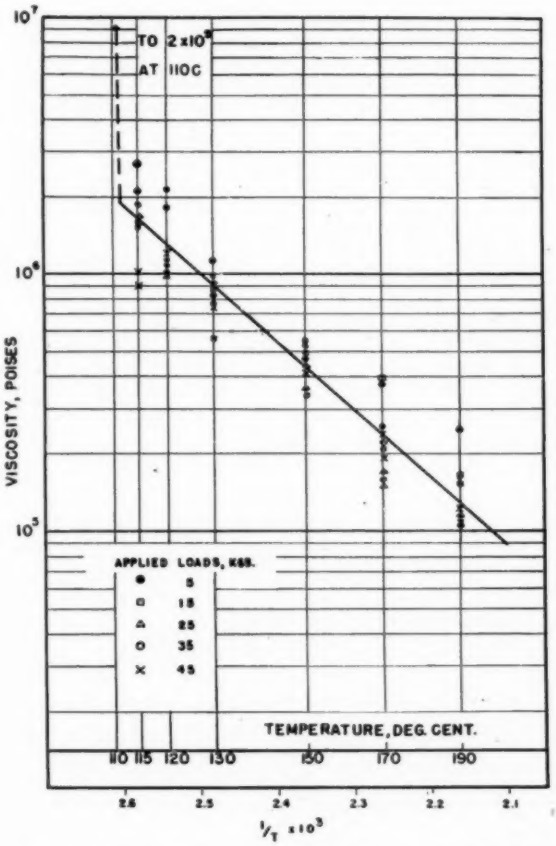


FIG. 8. Viscosity-temperature relation for polyethylene.

In the evaluation of this integral, two cases are treated.

Case 1: The viscous liquid completely fills the space between the plates so that $R=a$ at all times, where a is the radius of the plates. In this case, Eq. (11) on integration becomes:[4]

$$F=-\frac{3\pi\eta a^4}{2h^3}\frac{dh}{dt}. \tag{12}$$

Equation (12) integrated with respect to t gives

$$\frac{1}{h^2}=\frac{4F}{3\pi\eta a^4}t+C_1, \tag{13}$$

where

C_1 = constant of integration.

Equations (12) and (13) are known as Stefan's equations.[2]

Case 2: The viscous material does not completely fill the space between the plates. Thus, the radial distance, R (radius of the specimen), changes with time, due to the motion of the upper plate. In this case, Eq. (11) is integrated from $r=0$ to $r=R$ where R now is a function of the time, t. This can be done by assuming that the material between the plates is always essentially in the form of a cylindrical sample. It has already been assumed that the viscous material tested is incompressible so that the volume, V, is a constant. In this case, integration of Eq. (11) leads to

$$F=-\frac{3\eta V^2}{2\pi h^5}\frac{dh}{dt}. \tag{14}$$

Equation (14), on integration with respect to t, gives

$$\frac{1}{h^4}=\frac{8\pi F}{3\eta V^2}t+C_2, \tag{15}$$

where

C_2 = constant of integration.

The values of the constants of integration, C_1 of Eq. (13) and C_2 of Eq. (15), respectively, are the values of $1/h^2$ and $1/h^4$ at $t=0$, subject to the assumption that Stefan's equation is obeyed at $t=0$, and therefore do not correspond to the actual value of h at $t=0$. They are of no importance in the determination of the viscosity, η, from Eqs. (13) and (15), since η is calculable from the slope of the $1/h^2$ or $1/h^4$ *vs.* time curve.

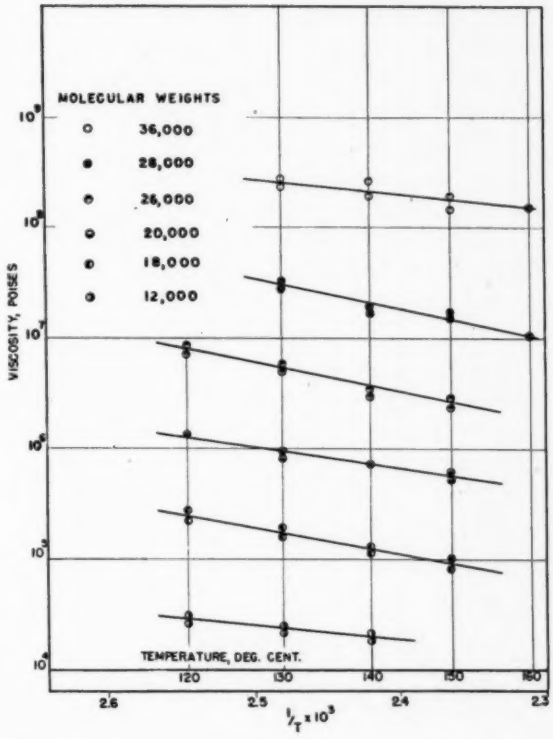


FIG. 9. Viscosity-temperature characteristics of various molecular weight grades of polyethylene.

Equations (13) and (15) describe the flow between two parallel plates moving toward each other in terms of the plate separation, h, the time, t, and the viscosity, η.

Calculation of Viscosity from Plate Separation and Time Measurements

In any parallel plate plastometer, a sample is compressed between the two plates, and the plate separation, h, is determined as a function of time. Two types of parallel plate plastometers may be considered corresponding to the two equations, Eqs. (13) and (15).

Type 1: The plates have the same area as the ends of the specimen. In this case, the space between the plates is completely filled with the viscous material at all times, and Eq. (13) is directly applicable in its integrated form. According to this equation, a plot of $1/h^2$ *versus* t will

[4] *The Mechanical Properties of Fluids.* A collective work. Applied Physics Series. (D. Van Nostrand Company, New York, 1924), pp. 113–119.

be linear with the slope of the straight line depending inversely on the viscosity. An example of the use of this type will be given later.

Type 2: The plates are large compared to the area of the ends of the specimen. In this case, the radius of the contact area, R, varies with the plate separation, h, and Eq. (15) is applicable. According to this equation, a plot of $1/h^4$ *versus* t will be linear with the slope of the straight line depending inversely on the viscosity.

The parallel plate plastometer in this laboratory may be used as either Type *1* or *2*. As a matter of experimental convenience, it has been operated almost exclusively as Type *2*, and Eq. (15) has been used to calculate the viscosity. For this case, the viscosity may be expressed from Eq. (15) in terms of the applied load; slope, m, of the $1/h^4$ *vs.* time curve; and specimen volume as follows:

$$\eta = 8.21(10)^6(W/mV^2). \qquad (16)$$

In this equation, the applied force, F, previously expressed in dynes, was replaced by W expressed in kilograms, m is expressed in cm^{-4} $sec.^{-1}$, V in cm^3, and η in poises.

If the experimental data obtained on a material follow the $1/h^4$ *versus* t straight line law, then the viscosity of the material can be calculated from Eq. (16). Accordingly, the linearity of the $1/h^4$ *versus* t plot was adopted in this work as the criterion for the measurement of predominantly viscous deformation. Other workers in this field, namely V. E. Meharg, L. E. Welch, and W. A. Miller;[5] R. L. Peek;[6] and N. M. Foote,[7] have used the differential form as given by Eq. (14) for the calculation of viscosities. R. W. Griffiths,[8] in a paper on the plasticity of rubber, considered the use of Eq. (15) for the calculation of viscosity (a simplified theoretical treatment is given by Healey in the appendix), and showed a $1/h^4$ *vs.* time curve for rubber. The possibility of using the linearity of the $1/h^4$ *vs.* time plot for separating the viscous deformation from other effects was apparently not investigated further.

It must be pointed out that this linear relation

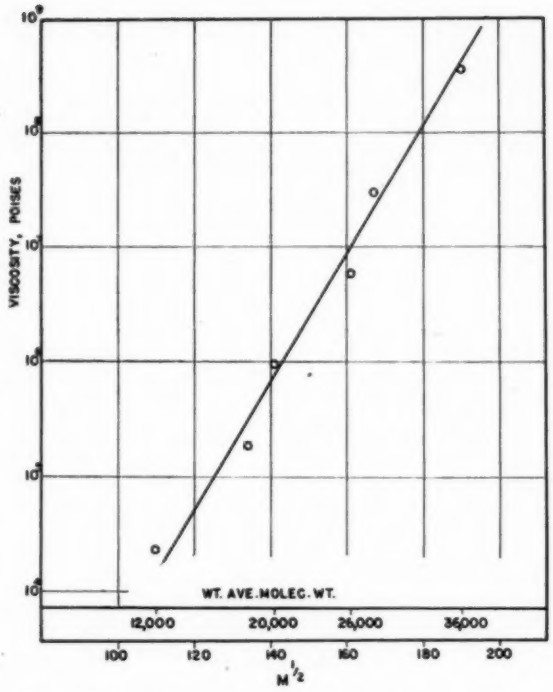


FIG. 10. Viscosity of polyethylene resins as a function of the square root of the weight average molecular weight. Temperature = 130°C.

was derived for Newtonian viscous flow. Thus, the fact that a given material obeys this linear relation means that the material behaves as a Newtonian fluid at the rates of shear employed in the test. Mathematically, Dillon and Johnston[9] were able to obtain approximate expressions for the average rate of shear in different types of plastometers. They showed that the average rate of shear for parallel plate plastometers is in the neighborhood of 0.1 $sec.^{-1}$ (extrusion plastometers may be run at rates of shear in the neighborhood of 500–1000 $sec.^{-1}$). The parallel plate plastometer, therefore, measures viscosity at very low rates of shear. Non-Newtonian viscous flow is characterized by non-linearity of the rate of shear *versus* shear stress curve. These curves, however, may appear to be approximately linear for some non-Newtonian materials over the relatively small range and at the low rates of shear available in this instrument. Thus, it is impossible to determine, even by the tenfold variation of load available for the plastometer in this laboratory, the shape of the rate of shear

[5] V. E. Meharg, L. E. Welch, and W. A. Miller, unpublished data.
[6] R. L. Peek, Jr., J. Rheology **3**, 53, 345 (1932).
[7] N. M. Foote, Ind. Eng. Chem. **36**, 244 (1944).
[8] R. W. Griffiths, Trans. I. R. I. **1**, 308 (1925–26).
[9] J. H. Dillon and N. Johnston, Physics **4**, 225 (1933).

vs. shearing stress curve at other than very low rates of shear. Hence, the Newtonian or non-Newtonian character of the material remains indeterminate in parallel plate plastometer measurements. This instrument, combined with the mathematical technique developed, measures the absolute viscosity of very viscous high polymers in the range 10^4 to 10^9 poises, subject, of course, to the limitation noted above.

EXPERIMENTAL RESULTS ON SOME THERMOPLASTICS

Polyethylene

The viscosity of a polyethylene resin compound identified as DE-3401 Natural (polyethylene resin stabilized with 0.2 percent Akroflex C) was measured over an 80-degree temperature range at different applied loads.

Typical $1/h^4$ *vs.* time curves are shown in Figs. 6 and 7 at 115°C and 190°C, respectively. These curves were obtained after a thirty-minute preheat so that temperature equilibrium was established prior to load application. These graphs show that the $1/h^4$ *vs.* t curves are not linear until 300 to 1000 seconds have elapsed, with this period of non-linearity decreasing with increasing temperature. During this initial period of non-linearity, the $1/h^4$ *vs.* t curves are convex upward.

In order to explain the initial non-linear portions of these curves, two effects have to be considered. In the first place, the theory is not applicable in the very early stages of a test, since the experimental conditions do not satisfy the mathematical boundary conditions assumed in the derivation of Eqs. (13) and (15). However, after the first few seconds, non-linearity cannot be ascribed to this situation, but is due to the fact that the deformation is not simply viscous in character. Appreciable "delayed elastic" effects may be present which mask the viscous deformation. From the character of the three regions of a deformation-time curve, it is seen that the viscous deformation which is non-recoverable continues at a constant rate after the sum of the "elastic" and "delayed elastic" deformation has become constant (or at least negligibly small). Thus, it is expected that the criterion of viscous flow, namely the linearity of a $1/h^4$ *vs.* t plot, will be obeyed except in the initial period when the "elastic" deformations are predominant. This is the experimental observation as represented by the curves of Figs. 6 and 7. After an initial period of non-linearity, the $1/h^4$ *vs.* time curves become straight lines, and these latter represent viscous flow. The decrease in time of the initial portion of non-linearity with increasing temperature is also in agreement with what is known about the temperature dependence of "delayed elasticity."[10]

The viscosity of the material is calculated from the slope of the linear portion of the $1/h^4$ *vs.* t curves by the use of Eq. (16).

The experimental results will be presented in the form of $\log \eta$ *vs.* $1/T$ curves. This plot will give a straight line if these materials follow the well-known viscosity-temperature law,

$$\eta = Ae^{(B/T)}, \tag{17}$$

where A and B = material constants, η = viscosity in poises, and T = absolute temperature. Activation energies for viscous flow, which are proportional to the constant B, may be compared if the slopes of the $\log \eta$ *vs.* $1/T$ curves are determined.

The experimental results on polyethylene (DE-3401) are shown in Fig. 8. This figure contains all the points of duplicate tests at five different loads (5, 15, 25, 35, and 45 kilograms) at temperatures of 110°, 115°, 120°, 130°, 150°, 170°, and 190°C. From the erratic spread of the experimental points at different loads, it is concluded that, although a change in load undoubtedly has some effect on the viscosity, no definite trend dependent on load was determined.

At 110°C the viscosity of polyethylene is about a thousand times as high as at 115°C, with the only reliable experiment at 110°C obtained under a 5 kg load (at higher loads the specimen was partially crushed). This result is in agreement with the known behavior of polyethylene which shows a phase transition[11] in the vicinity of this temperature.

The $\log \eta$ *vs.* $1/T$ plot is a straight line from 115°C to 190°C. It is to be noted that the variation of viscosity with temperature for poly-

[10] R. F. Tucket, Trans. Faraday Soc. **39**, 158 (1943).
[11] R. F. Clash, Jr. and L. M. Rynkiewicz, Ind. Eng. Chem. **36**, 279 (1944).

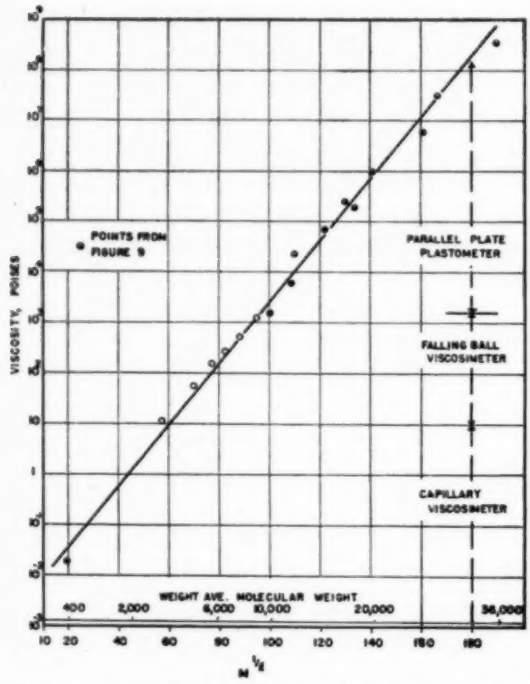


FIG. 11. Viscosity of the polyethylene resins and paraffin wax-polyethylene resin mixtures as a function of the square root of the weight average molecular weight. Temperature = 130°C.

ethylene is relatively small, roughly a tenfold change taking place over a range of 80 degrees, indicating a low activation energy for viscous flow.

The effect of preheating the specimen was investigated, and it was found that preheating had no appreciable effect on the measured viscosity. If the sample is not preheated, the initial portion of non-linearity of the $1/h^4$ *vs.* t curve may be somewhat longer, but the slope of the straight line portion is not affected. Accordingly, all further experiments were run without preheat as a matter of convenience.

The viscosity of polyethylene compounds (polyethylene resin stabilized with 0.2 percent Akroflex C) of various weight average molecular weights was also determined. Six polyethylenes of weight average molecular weight in the range 12,000 to 36,000 were available.

The weight average molecular weight (relative) of these materials was obtained by the Staudinger method as modified by Kemp and Peters.[12] The determination, made at 130°C in tetralin, was based on the equation

$$M=\frac{K_{cm}\log_{10}\eta_r}{C}, \tag{18}$$

where

M = weight average molecular weight,
η_r = relative viscosity,
C = base molar concentration of polymer in tetralin at temperature of determination (130°C),
= 4.0 g per liter of solution at 130°C/14, and
$K_{cm}=4.03\times10^4$ g per liter of solution at 130°C.

The experimental values of M used in this paper were taken from data obtained by Mr. C. S. Myers.[13]

The viscosity was measured for each material over a convenient temperature range. The results are given in the form of $\log\eta$ *vs.* $1/T$ curves in Fig. 9. Linear behavior is observed for all of these materials over the temperature range investigated. All of the curves have about the same slope, indicating that the activation energy for viscous flow does not change appreciably over this range of average molecular weights.

These data offer a test of the applicability of Flory's relation[14] to polyethylene. The parallel plate viscosity for these resins may be identified with the "melt viscosity." It was shown by Flory that the melt viscosity of polymerized linear polyesters varies with the weight average molecular weight and the temperature according to the equation

$$\log\eta=D+\frac{B}{T}+CM^{\frac{1}{2}}, \tag{19}$$

where η = melt viscosity, T = absolute temperature, M = weight average molecular weight (relative), and D, B, and C = constants.

TABLE I.

Weight average molecular weight	Approximate time (seconds) to establish linearity of $1/h^4$ *vs.* t at 130°C
12,000	120
18,000	150
20,000	350
26,000	400
28,000	700
36,000	1000

[12] A. R. Kemp and H. Peters, Ind. Eng. Chem. **35**, 1108–1112 (1943).

[13] C. S. Myers, unpublished data.

[14] P. J. Flory, J. Am. Chem. Soc. **62**, 1057 (1940).

Figure 10 is a plot of $\log \eta$ *versus* the square root of the weight average molecular weight at 130°C. A straight line is obtained confirming Flory's relation for polyethylene.

Since the slopes of the $\log \eta$ *vs.* $1/T$ curves are essentially independent of M (Fig. 9), the constants (D, B, and C) of Eq. (19) can be determined. This gives

$$\log \eta = -6.14 + \frac{2015}{T} + 0.05 M^{\frac{1}{2}}, \qquad (20)$$

which was found to agree approximately with all the experimental results.

From these relations, it is evident that the parallel plate viscosity can be used with advantage in estimating weight average molecular weights. The use of the parallel plate plastometer extends considerably the range of viscosity that may be measured conveniently, thereby expanding the range of molecular weights that can be estimated.

It is interesting to consider the variation of "delayed elasticity" as a function of the average molecular weight. The length of time necessary to reach the linear portion of the $1/h^4$ *vs.* t curve may be considered an index of "delayed elasticity." The length of this time interval was estimated, at 130°C, from the $1/h^4$ *vs.* t curves. The data given in Table I were obtained.

TABLE II.

Compound B-660	
Vinyl chloride-acetate copolymer, grade VYNS	82.20
Vinyl chloride-acetate copolymer, grade VYHH	15.00
Lubricant, stabilizer, etc.	2.80
	100.00
Compound B-661	
Vinyl chloride-acetate copolymer, grade VYHH	97.35
Lubricant, stabilizer, etc.	2.65
	100.00

These data show that "delayed elasticity" increases markedly with increasing average molecular weight. It is to be noted that the measurements on these materials were made without any preheat. Since it requires about two minutes to obtain temperature equilibrium in the specimen, the two low molecular weight compounds show a "delayed elasticity" of very short duration at this temperature. The trend with increasing weight average molecular weight, however, is unquestionable.

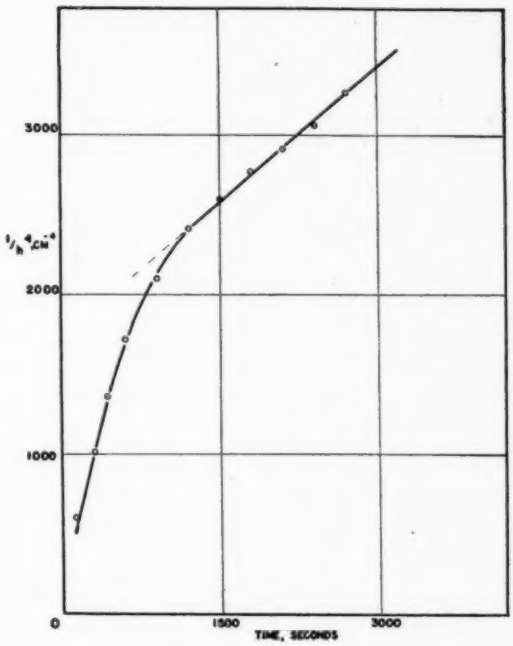


FIG. 12. Typical $1/h^4$ *vs.* time curve for a vinyl chloride-acetate resin compound (B 660) at 155°C.

Paraffin Wax—Polyethylene Mixtures

Parallel plate viscosity measurements were carried out on a set of paraffin wax-polyethylene resin mixtures in the 50 percent to 100 percent resin content range. These compounds were stabilized mixtures of a refined low melting, household grade of paraffin wax ("Esso Wax," m.p. 52°C) and a standard polyethylene resin (Grade DYNH, wt. av. mol. wt. 20,000) containing 0.1 to 0.2 percent antioxidant.

Viscosity figures for the mixtures of 0 percent to 50 percent resin content were made available by Mr. C. S. Myers.[13] The viscosity of the paraffin wax was determined by a Ubbelohde Viscosimeter (Type IA); the viscosities of the mixtures (up to 50 percent resin content) were obtained by a falling ball viscosimeter. Weight average molecular weights, obtained by the method of Kemp and Peters as already described, were taken from Mr. Myers' data.

Application of Flory's equation (Eq. (19)) to these viscosity data at 130°C showed a remark-

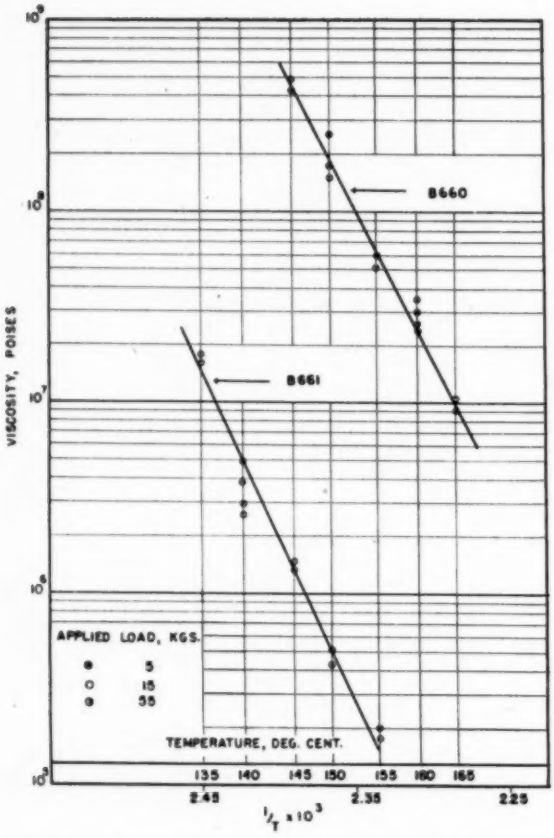


FIG. 13. Viscosity-temperature characteristics for two vinyl chloride-acetate resin compounds.

able degree of conformity. Figure 11 is a plot of log η *versus* the square root of the weight average molecular weight at 130°C. This figure contains all the data for the mixtures, paraffin wax alone, and all the polyethylenes of various molecular weights from Fig. 10. It will be noted that very good agreement is obtained among three different methods of viscosity measurement with the viscosity varying from 10^{-2} to about 10^{9} poises. It is also observed that Flory's relation evidently applies over a wide range of weight average molecular weights (400 to 36,000).

Vinyl Chloride-Acetate Resin Compounds

The theory and the discussion of the results for polyethylene apply with minor modifications to vinyl chloride-acetate resin compounds. Viscosity measurements were made on two unplasticized vinyl chloride-acetate resin compounds (composition in percent by weight) (see Table II).

A typical $1/h^4$ *vs.* time curve for compound B-660 (at 155°C) is shown in Fig. 12. This curve is very similar to the curves for polyethylene, except that the "delayed elastic" region is much more pronounced. The straight line portion is not reached for about 25 minutes. Similar behavior was observed with compound B-661. These $1/h^4$ *vs.* t curves must be interpreted with caution due to these long initial periods of non-linearity. For the two compounds above, very good straight lines were obtained over the temperature range investigated. Viscosities were calculated from the slope of the straight line portions.

The viscosities follow the usual log η *vs.* $1/T$ straight line relation, as illustrated in Fig. 13. This figure also shows that the effect of changing the applied load is negligible. It is to be noted that the viscosity variation with temperature is much higher than for polyethylene, indicating a much higher activation energy for viscous flow.

This technique of viscosity measurement has

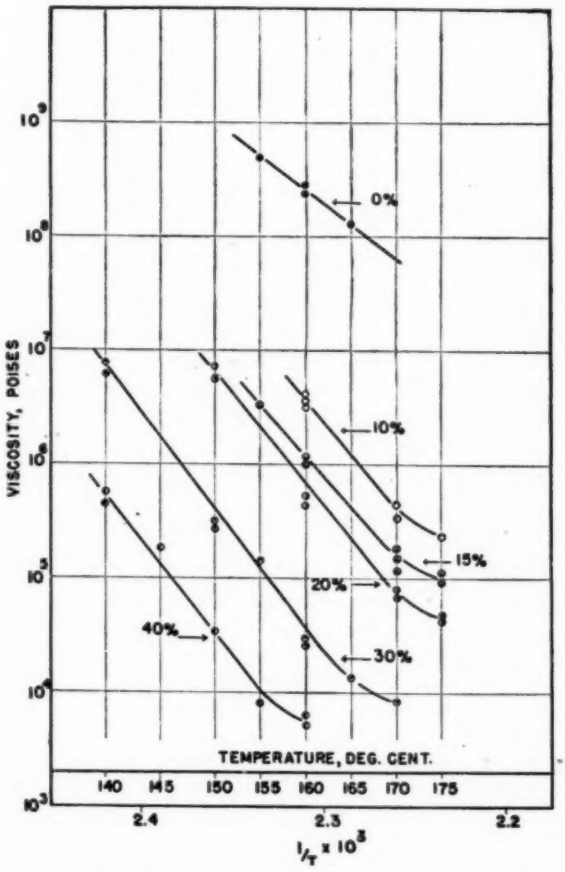


FIG. 14. Viscosity-temperature characteristics of VYNS+dioctyl phthalate systems.

also been applied successfully to plasticized vinyl chloride-acetate resin compounds. The general formula used for these compounds was (percent by weight):

Vinyl chloride-acetate resin *VYNS*	97-X
Dioctyl phthalate	X
Stabilizers and lubricants	3
	100

The log η *vs.* $1/T$ curves at different plasticizer concentrations are shown in Fig. 14. It will be noted that, for the plasticized compounds, straight lines are obtained up to a certain temperature, above which the viscosity does not decrease as rapidly with increasing temperature as predicted by the straight line. It is also noteworthy that the slope of the log η *vs.* $1/T$ curves, that is, the activation energy for viscous flow, is essentially unchanged by the addition of plasticizer.

Figure 15 is a plot of log η *vs.* percent dioctyl phthalate at 160°C. This figure shows that very large changes in viscosity can be accomplished by varying the plasticizer content. The relation is linear, except in the 0 to 10 percent plasticizer range, pointing to a close parallel between the effects of increasing temperature and increasing plasticizer concentration.

The parallel plate plastometer has also been operated as Type *1*; that is, the plates have the same area as the ends of the specimen. In this case, the sample has to be accurately lined up with the plates. For this test a $1/h^2$ *vs.* time straight line law is applicable (Eq. (13)). Experiments on polyethylene and vinyl chloride-acetate resins showed that this measurement of viscosity agrees with that obtained with the Type *2* plastometer. Experimentally, it has been more convenient to operate the plastometer as Type *2* because the alignment of the specimen is less critical.

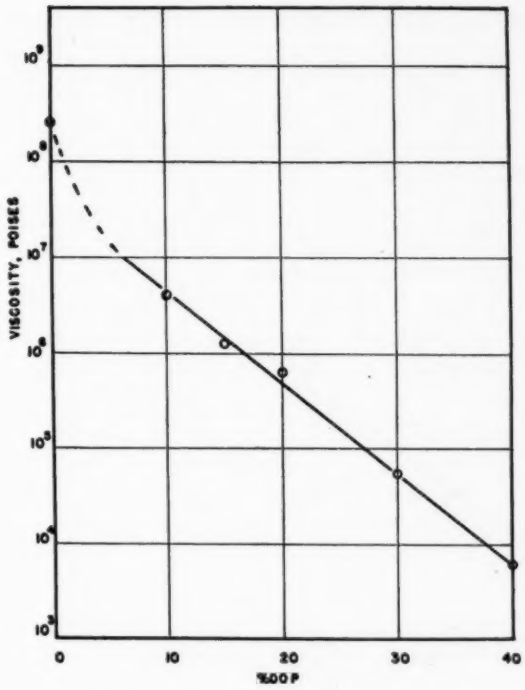


FIG. 15. Viscosity-plasticizer content curve for VYNS+dioctyl phthalate at 160°C.

ACKNOWLEDGMENT

The parallel plate plastometer used in this work was designed by Messrs. H. A. Swallow and W. R. Yaroshuk, and constructed by Tinius-Olsen Testing Machine Company. Grateful acknowledgment is also made to our associates in the Physics Division of the Bakelite Development Laboratories, particularly to R. F. Clash, Jr. and A. P. Wangsgard, for many helpful discussions on the various phases of this work. It is also a pleasure to acknowledge the helpful assistance derived from the earlier unpublished work on the parallel plate plastometer of Messrs. V. E. Meharg, L. E. Welch, and W. A. Miller of our Research Laboratories.

The General Relations for Flow in Solids and Their Application to the Plastic Behavior of Tire Cords

W. James Lyons*
*Southern Regional Research Laboratory,** New Orleans, Louisiana*
(Received January 30, 1946)

Numerous researches on metals, rocks, glass, rubber, and textiles are cited to show that what may be regarded as *normal* creep in amorphous and polycrystalline solids conforms to the general relation

$$\epsilon=\epsilon_1+at+b\log t,$$

where ϵ is the total strain at time t, ϵ_1 is a parameter interpretable as the approximate initial strain, and a, b are other parameters. A special form of this equation, having $a=0$, has been frequently applied. Reference is made to several other studies on the same class of solids, which have established for normal relaxation the relation

$$\sigma=\sigma_1-\beta\log t,$$

where σ is the stress at time t, σ_1 is the stress at unit time, and β is a parameter. A theoretical foundation for both equations is provided by the reaction-rate theory of plastic flow. Observations in the present study indicate that creep extension in cotton and rayon tire cords over prolonged periods of time follows the above creep equation, and, in general, is not adequately represented by an equation omitting the term at. From the presence of this term in the equation, the existence of a component of the viscous type in tire cord growth is deduced. The pattern of creep recovery in cotton tire cord appears to be set by the behavior of the cord in creep extension.

THE phenomena of creep and relaxation arising from flow in solids has been studied in a wide variety of materials ranging from soft bread doughs to the rocks in the crust of the earth. Attention has frequently been called to the technical importance of these flow phenomena in rubber products, textiles (particularly the industrial textiles, such as tire cord), high polymeric plastics, metals, and numerous other materials normally accepted as "solids." Considering the great diversity of chemical composition, molecular character, and general physical behavior, it is indeed remarkable that the laws describing the general pattern of creep and relaxation in a large number of these substances take the same two analytical forms. The only characteristic which is evidently common to all of this broad group of substances is a structure which is predominantly either amorphous or polycrystalline. This observation suggests that the types of behavior in creep and relaxation common among these substances are to be associated with randomness of structure.

Any attempt to bring experimental data on a wide variety of otherwise unrelated materials into conformity with a general law should proceed with caution. On the other hand, the scientific ideal of systematized knowledge dictates that such possible generalizations be recognized, though they should not become the means for the entrenchment of misinformation. In the field of plastic-solid flow, a dispersive trend has prevailed, new empirical equations having been advanced with no apparent attempt having been made to correlate the data with earlier, applicable equations. The benefit of a *general*, empirical law, especially if there is a concomitant explanatory theory, lies in its suggestion of the elementary processes giving rise to exceptional behavior in particular cases.

THE LAW OF CREEP

The deformation resulting from creep under constant load at constant temperature in a large number of materials is in conformity with the relation

$$\epsilon=\epsilon_1+at+b\log t, \qquad (1a)$$

where ϵ is the strain at time t, ϵ_1 is a parameter having the dimensions of strain, and a, b are

* Present address: The Firestone Tire & Rubber Company, Akron 17, Ohio.
** One of the laboratories of the Bureau of Agricultural and Industrial Chemistry, Agricultural Research Administration, U. S. Department of Agriculture.

parameters depending on the load, material, and ambient experimental conditions. $\epsilon(t)$ in Eq. (1a) is to be regarded as a function describing the creep process; in general, some other function of time describes the deformation in the initial stage. Thus, while the two processes are continuous physically, Eq. (1a) is invalid at $t=0$, and no attempt to evaluate $\epsilon(t)$ at $t=0$ should be made. The parameter ϵ_1 may be interpreted as an *extrapolated* initial strain, for if unity on the time scale used with Eq. (1a) is taken as the instant at which the initial strain is completed, and the insignificant contribution of at during the very brief initial stage is neglected, it will be seen that the deformation at later times behaves *as if* its initial value were ϵ_1. In general, ϵ_1 only approximates the strain actually occurring during the initial stage. Equation (1a) was advanced by Weaver[1] to represent the total creep extension of various steels at constant stress and temperature in tests lasting, in some cases, five years. The parameter corresponding to ϵ_1 was negative, however, and was not given the interpretation outlined above.

An approximate relation of a type similar to Eq. (1a) omits the term at, thus implicitly placing $a=0$, and giving

$$\epsilon=\epsilon_1+b\log t. \tag{1b}$$

The creep component of the total deformation is here represented solely by a logarithmic function of the time. The possibility of so representing creep appears to have been first recognized by Boltzmann[2] in 1874. He used an equation of the type of (1b) to express the torsional deformation in a filament as a function of time.

In order to define the function $\epsilon(t)$ so that it will be continuous at $t=0$, the argument of the logarithmic term may be written in the binomial form $(ct+1)$. Applying this modification, Eqs. (1a) and (1b) become

$$\epsilon=\epsilon_1+at+b\log(ct+1), \tag{1c}$$

and

$$\epsilon=\epsilon_1+b\log(ct+1), \tag{1d}$$

where c is a parameter having the dimension of inverse time. For moderately large values of ct, however, the relation

$$b\log(ct+1)\approx b\log c+b\log t$$

holds very closely. Hence, if ϵ_1 in Eqs. (1a) and (1b) is assumed to include the constant term $b\log c$, no distinction between Eqs. (1a) and (1c) or (1b) and (1d) can be made on the basis of conformity with experiments in which $ct\gg 1$. For this reason, the form $\log(ct+1)$ appears to be more comprehensive than is usually necessary for the representation of creep. Equation (1c), which may be regarded as the most general law of creep under constant load, has not been found in the literature. Equation (1d), however, has been used, as indicated below.

Since its application to creep in steel by Weaver,[1] Eq. (1a) has been employed to represent the slow yield of limestone, and of a single crystal of halite under compression,[3] as well as the creep in molded phenolic plastics.[4] Other applications of this equation have been cited by Findley.[5] Furthermore, there are in the literature several researches on creep which have yielded results evidently in conformity with Eq. (1a); the reports, however, do not cite the equation. Among these may be mentioned the tensile tests of Trouton and Rankine[6] and Andrade[7] on lead, those of Bailey[8] on molybdenum steel at elevated temperatures, and the three-hour compressive tests of Lyons[9] on copper and lead. The data published on creep under tension in eight organic plastics of various types[10–12] appear to be in accord with Eq. (1a) when moderate deforming stresses are employed. Examination of the results of Findley[5,12] indicates that under high stresses the creep in cellulose acetate proceeds as if a in Eq. (1a) had a negative value, which is a rather unacceptable condition. Griggs[3] has noted a

[1] S. H. Weaver, Trans. A.S.M.E. **58**, 745 (1936).
[2] L. Boltzmann, Pogg. Ann. **7**, 624 (1876), Sitz. K. Akad. Wiss. Wien, Math.-Naturwiss. Classe **70**, 275 (1874).
[3] D. Griggs, J. Geol. **47**, 225 (1939).
[4] D. Telfair, T. S. Carswell, and H. K. Nason, Mod. Plastics **21**, 137 (Feb. 1944).
[5] W. N. Findley, *Symposium on Plastics* (Philadelphia, A.S.T.M., 1944), p. 118.
[6] F. T. Trouton and A. O. Rankine, Phil. Mag. [6] **8**, 538 (1904).
[7] E. N. Andrade, Proc. Roy. Soc. **A84**, 1 (1910); **A90**, 329 (1914).
[8] R. W. Bailey, J. App. Mech. (Trans. A.S.M.E.) **58**, (A)1 (1936).
[9] W. J. Lyons, J. App. Phys. **9**, 641 (1938).
[10] B. Chasman, Mod. Plastics **21**, 145 (Feb. 1944).
[11] J. Delmonte and W. Dewar, Mod. Plastics **19**, 73 (Oct. 1941).
[12] W. N. Findley, Mod. Plastics **19**, 71 (Aug. 1942).

similar behavior in creep in shale. The physical interpretation of Eq. (1a), however, provides an explanation of this behavior at high stress.

In the field of textiles, several studies have yielded results which are amenable to expression by means of Eq. (1a). Among these are the long duration creep results of Leaderman[13] on silk, and viscose and acetate rayon filaments. The data on single Nylon filaments, however, present an anomaly with reference to Eq. (1a) which resembles that found in cellulose acetate under high stress. The creep results of Busse and co-workers[14] on rayon tire cord suggest that both the parameters a and b increase with temperature. In the study of Dillon and Prettyman,[15] creep in cotton and rayon cords, at three different temperatures, is represented by straight lines on an *elongation vs. logarithmic-time* graph, indicating a value of zero for a. However, the curves found by Dillon and Prettyman for Nylon cord show a definite upturn at later times. In the experimental portion of the present study, results on tire cord creep extending into thousands of minutes were obtained. For these results Eq. (1a), as will be shown, provides a good representation.

The creep of soft copper[16] under constant tensile load, as well as that of viscose yarns,[17] has been shown to be in accord with Eq. (1d), the coefficients b and c being functions of certain material characteristics and experimental conditions. The omission of the constant rate term at from the equation, as made in Eqs. (1b) and (1d), appears to be applicable only in short duration tests, in which the viscous type of flow represented by at has little opportunity to become appreciable. Equation (1b) has been used to represent tensile creep in rubber, glass, and annealed copper, platinum, silver, and gold wires.[18] It has also been applied to creep results on cotton and rayon fibers in tension[19] and lead cylinders in compression[20] in thirty-minute tests. That creep in soft iron, also, is not a direct function of the time, but a function of the logarithm of the time, has been noted.[21]

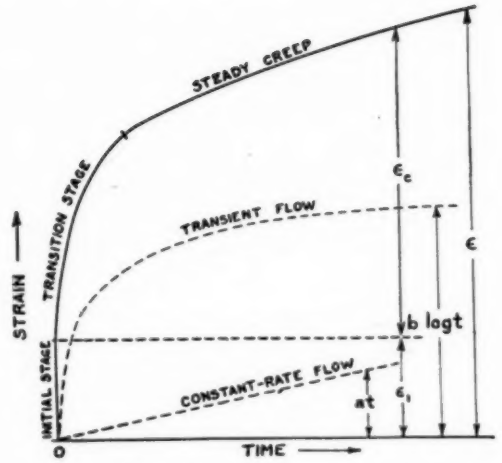


FIG. 1. Hypothetical diagram of the components of total creep deformation, as applied to terms in Eq. (1a).

THE LAW OF RELAXATION

The relaxation of stress at constant deformation also has been studied in a variety of materials. The general, empirical relation, for behavior at constant temperature, may be written

$$\sigma = \sigma_1 - \beta \log t, \tag{2a}$$

where σ is the stress at time t, σ_1 is a parameter interpretable as the initial stress, with the same limitations as were put on the interpretation of ϵ_1, and β is a parameter depending on the initial stress, the material, and the particular temperature of the experiment. A variant of Eq. (2a) has been announced, and we give it forthwith:

$$\sigma = \sigma_1 - \beta \log (\gamma t + 1), \tag{2b}$$

where γ is a parameter like c.

Equation (2b), which, like Eq. (1c), is continuous at $t=0$, was given in 1904 by Trouton and Rankine[6] for the decay of longitudinal stress in lead wire. In the same article they indicate that Eq. (2a) fits the data on the relaxation of torsional stress in a lead wire. In subsequent research, Eq. (2b) was applied to stress relaxation under constant strain in concentrated gels of

[13] H. Leaderman, *Elastic and Creep Properties of Filamentous Materials and Other High Polymers* (Washington, Textile Foundation, 1943).
[14] Busse, Lessig, Loughborough, and Larrick, J. App. Phys. **13**, 715 (1942).
[15] J. H. Dillon and I. B. Prettyman, J. App. Phys. **16**, 159 (1945).
[16] A. Nadai and E. A. Davis, J. App. Mech. (Trans. A.S.M.E.) **58**, (A)7 (1936).
[17] J. J. Press, J. App. Phys. **14**, 224 (1943).
[18] P. Phillips, Proc. Phys. Soc. London **19**, 491 (1905).
[19] R. L. Steinberger, Textile Research **6**, 325 (1936).
[20] W. J. Lyons, J. App. Phys. **10**, 651 (1939).
[21] H. Cassebaum, Ann. d. Physik **34**, 106 (1911).

gelatine,[22] and recently it has been applied to results on rayon yarn.[17] It has been observed that Eq. (2a) represents the decay of stress in rubber under tension.[18] The graphical plots of relaxing stresses in rubber against *log t* yield straight lines.[23] Equation (2a) has also been found to hold for relaxation in cotton, rayon, and Nylon tire cords at various temperatures.[24]

No equation for relaxation containing a viscous-flow term such as *at* in Eq. (1a) has been found in the literature. There are some data, such as those of Trumpler[25] on relaxation at constant elongation in steel, which seem to call for such a term (with a negative sign), but in such cases the major portion of the relaxation is satisfactorily represented by Eq. (2a).

THEORETICAL CONSIDERATIONS

At least four equations[5, 7, 26, 27] which have been advanced to relate stress σ, strain ϵ, and time t during flow in various solids involve power functions of t combined with coefficients having incomprehensible dimensions. In general, the power is not integral. In many of the applications, with a suitable choice of coefficients and exponents, good fits to experimental data have been made. The utility of the equations, however, is limited to convenient analytical expression, for purposes of extrapolation or prediction. Even aside from the difficulty of interpreting fractional powers of t, these empirical equations provide no insight into the physical processes occurring during plastic deformation. Equations (1a–d) and (2a, b), on the other hand, are not only interpretable in terms of gross physical effects, but find explanation in a theory of molecular mechanism.

As long ago as 1904 it was recognized by Trouton and Rankine[6] that the plastic deformation of solids involves a component of the elastic type and one of the viscous type. The former consists of flow which gradually decreases with time, and is more or less recoverable on removal of the deforming load. It has been designated by a seemingly endless number of names, among which "transient flow," used here, is one. The relation of these components (represented by dashed curves) to the total deformation ϵ (the solid curve) during creep, is shown in Fig. 1. Such a graphical analysis may be made without reference to a specific equation. Figure 1, however, lends itself well to the identification of the terms in Eq. (1a). The term ϵ_1 represents the deformation occurring in the initial stage, though as indicated above, numerically the representation is merely approximate. In the second term, a is the constant rate of flow characterizing the viscous component, which persists indefinitely. The transient flow is expressed with good fidelity by the final term $b \log t$. As Eq. (1a) indicates, the total strain at any instant during steady creep is given by the simple arithmetic sum of the components. The strain due to creep alone is $\epsilon_c = at + b \log t$.

The processes outlined in Fig. 1 have been interpreted in terms of another well-known creep equation,[28] namely,

$$\epsilon = \epsilon_1 + \lambda t + \epsilon_\infty [1 - \exp(-t/\tau)], \qquad (3)$$

where λ is the constant rate of viscous flow, ϵ_∞ is the final, maximum strain due to transient flow, and τ is the Maxwellian relaxation time for the system. This equation was developed for a system of models of the type considered by Maxwell in the derivation of his classical equation for relaxation. Involving the evaluation of four coefficients, instead of three as does Eq. (1a), Eq. (3) is less manageable. Furthermore, as was found in the present study on tire cord, the fit to experimental data which is finally achieved is not as good as that which Eq. (1a) readily gives. Equations of the type of (3), however, have found favor, presumably because of their theoretical foundations, which until recently Eqs. (1) and (2) have lacked. Except insofar as the transient term $b \log t$ is an approximation to $\epsilon_\infty(1 - \exp(-t/\tau))$, Eqs. (1) and (2) are not derivable from Maxwell's equation. Integration of the latter equation for

[22] A. O. Rankine, Phil. Mag. [6] **11**, 447 (1906).
[23] A. Tobolsky and H. Eyring, J. Chem. Phys. **11**, 125 (1943).
[24] H. R. Wakeham, Southern Reg. Res. Lab., private communication.
[25] W. E. Trumpler, Jr., J. App. Phys. **12**, 248 (1941).
[26] A. A. Michelson, Proc. Nat. Acad. **3**, 319 (1917); **6**, 122 (1920).
[27] P. G. Nutting, J. Frank. Inst. **191**, 679 (1921); **235**, 513 (1943).
[28] J. M. Burgers, *First Report on Viscosity and Plasticity* (Committee of Academy of Sciences at Amsterdam, Amsterdam, 1939), second edition, p. 28.

constant deformation leads to

$$\sigma = \sigma_1 \exp(-t/\tau), \qquad (4)$$

as the expression for the relaxation of stress. This equation, however, has been notoriously poor in representing experimental data. While it may be in accord with data on a few materials, it fails for all those cases in which the corresponding empirical Eq. (2a) holds. Trouton and Rankine[6] were among the earliest to abandon Eq. (4) in favor of Eqs. (2a) or (2b). Since the Maxwell equation is empirically unsatisfactory in the simplest applications, there appears to be little reason to accredit the underlying theory, and attempt to interpret more complicated processes in terms of it, as use of Eq. (3) implicitly does.

It is on the transient flow that most theoretical interest in creep and relaxation has centered. It will be noted that it is in the term for this component that Eqs. (1a) and (3) differ. The transient flow reappears in relaxation processes and is represented in Eqs. (2a) and (2b) by $\beta \log t$, or $\beta \log (\gamma t+1)$. Some of the earliest analytical considerations of solid flow[7,18] associated the transient part with local rupture, rotation, or reorientation of the submicroscopic elements of the material. Transient flow, as the integration of localized movements, was distinguished from the shearing process in continuing viscous flow. The later ideas considered the transient flow to be imposed on the viscous. This view is in line with the diagram of Fig. 1.

While qualitative theory has retained this picture of the general character of transient flow, the justification for the empirical, logarithmic form has been lacking. It is possible to set down a general differential equation for the process of deformation in solids, from which the initial and transient terms in Eqs. (1) and (2) can be derived and correlated. The treatment, however, is a mathematical formalism, not unlike those accompanying application of Maxwell's equation; it contributes nothing toward establishing the logarithmic form as the unique expression for specific molecular processes. The descriptive, reaction-rate theory of Eyring[29] does provide a basis for the logarithmic form in a mechanism at the molecular level. This theory leads to the following relation between stress and time:

$$\tanh \frac{B\sigma}{2} = \tanh \frac{B\sigma_1}{2} \exp(-ABGt), \qquad (5)$$

where σ_1 is the initial stress, G is an elastic modulus, and A and B are parameters involving temperature among other factors. Tobolsky and Eyring[23] have shown that, applied to longitudinal creep under constant load, beyond the range in which the cross section is changing rapidly, Eq. (5) leads approximately to Eq. (1b). For stress relaxation at constant elongation they derive Eq. (2a).

The constant-rate term at in Eq. (1a) does not emerge from the reduction of Eq. (5), which refers to the integration of localized slipping processes, such as might occur in crystalline solids or high polymeric network-structures. These processes consist of recrystallization arising from the assumption of new equilibrium positions by displaced atoms, and the reformation of ruptured bonds between chain molecules. In the complete Eyring theory of solid flow, however, there appears an equation for the motion of the amorphous portion of linear-polymeric substances ("segment motion"). If this equation is integrated independently for constant stress, the resultant expression yields a strain which increases at a constant time-rate. The theory in its present form, however, does not employ the outright addition of this strain to that given by Eq. (5). It does not appear feasible to dismiss the constant-rate term at as an insignificant experimental deviation from the logarithmic law, for in some materials, as lead in creep-extension[6] measurements for instance, the deformation resulting from viscous-like flow is the most prominent component. It may be concluded that in the creep of the numerous solids which behave in accordance with Eq. (1a), there is an actual underlying constant-rate flow corresponding to the term at. The constant-rate, viscous-type flow may be interpreted as a slow general distorting motion within a connected, reticular system, or between such systems, through a matrix of unassociated, amorphous substance. This slow movement could conceivably take place whether the motive force were an external load, or internal restoring stresses in the networks.

[29] H. Eyring, J. Chem. Phys. **4**, 283 (1936).

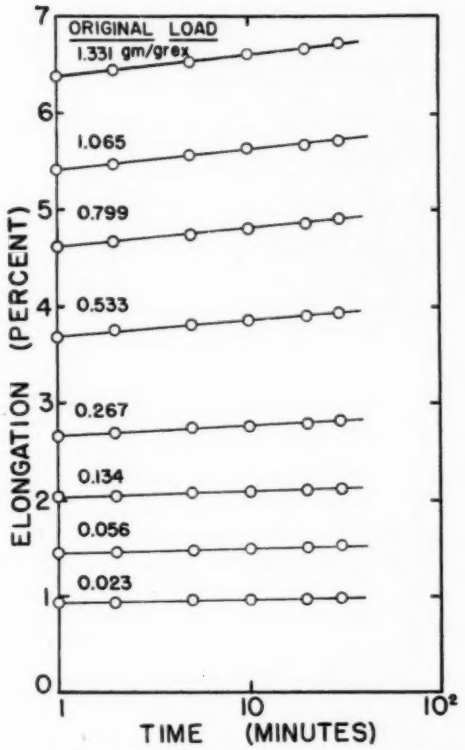


FIG. 2. Creep in a "hot-wet-stretched" 17/4/3 cotton cord.

As a measure of the constant rate of viscous flow in a material or product, the parameter a has technical importance. The comparison of values of a for different materials under equal loads will indicate which will offer the most resistance to continuous yielding or growth in service.

Reference was made above to the anomalous behavior of cellulose acetate strips and Nylon filaments with reference to Eq. (1a). A decelerated creep at the longer times was observed in these materials, contrary to behavior predicted by Eq. (1a). In the light of the foregoing analysis, the observed behavior indicates extensive reorientation of the chain molecules and enlargement of crystalline areas: an excessive strain-hardening effect. Such an effect might be expected in the linear high polymers, but would not be likely to be found in polycrystalline metals having a granular structure. Presumably, under the action of high stress, or because of easy susceptibility to orientation, the strain-hardening in the acetate and Nylon becomes so pronounced as to alter seriously the viscous properties of the material and depress the value of the parameter a. In the case of the cellulose acetate the strain over the anomalous region was not much below that at which fracture occurred. The strain-hardening in this region may be regarded as the preliminary stage of the rupture process, and abnormal effects are to be expected. This downward curvature of the *creep vs. log t* curve at high limiting values of t has a counterpart in the complete graph of Eq. (5).[23]

CREEP AND RECOVERY IN TIRE CORDS

The measurement in the laboratory of creep and other visco-elastic effects in tire cords provides a useful insight into the "growth" behavior of the cords in a tire in service. Creep experiments with a constant load would appear to parallel closely the conditions in a tire which lead to growth. The cords in an inflated tire on the wheel are continually under tension. The rubber encasing the cords contributes very little toward supporting the load on the tire, and limiting the plastic extension of the stressed cords. In this respect, conditions in a tire differ from those of stress-relaxation experiments of the type in which provision is made to hold the specimen at constant length.

Since the experimental phase of the present study was intended primarily as an exploration for constant-rate flow in tire cord, relaxation experiments have not been included. Relaxation experiments should, however, provide valuable data complementary to those on creep, when the plastic behavior of a particular cord is being exhaustively examined. In tensile creep, the decay of stress is accompanied by an increase in strain tending to offset the decay. By eliminating strain as a variable, the experiments at constant length provide a means for studying the stress history of a cord directly. The observed stress-relaxation in a cord is presumably a reflection of the adjustments taking place at the micellar and molecular levels, without the complication of concomitant deformation in the fiber substance. Furthermore, the glide of fiber on fiber, which is conceivable as a component of the observed creep in cords composed of staple fibers, is substantially eliminated in stress-relaxation experiments.

Creep measurements under constant load were made at room temperature on cotton tire cords

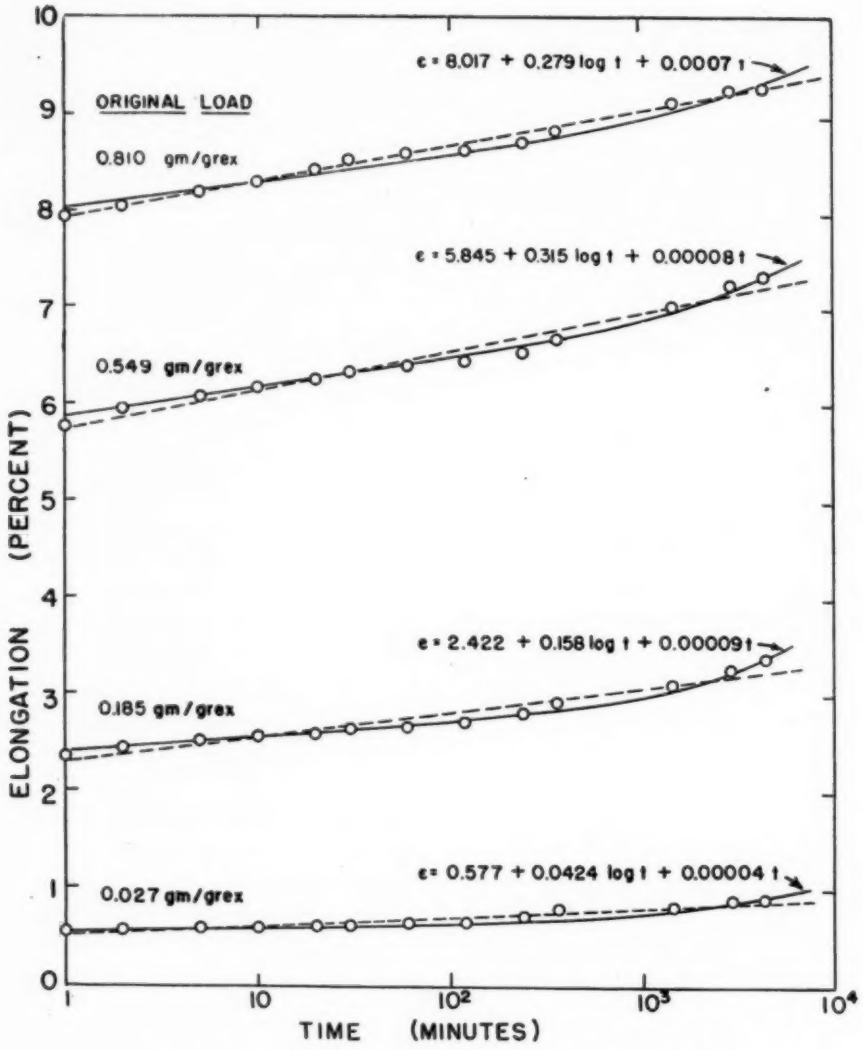


FIG. 3. Creep in a regular, wet-twisted 29/5/3 cotton cord.

of three different types and on a rayon cord. In general, for each cord sample at each of the several loads employed, the time-sequence of observations was made on five specimens. The mean values of the lengths of the specimens at selected times were used in plotting the *elongation vs. time* curves. The observations on the cords, which were hung in a vertical position, were made by means of a cathetometer. The experiments were conducted in an air-conditioned laboratory in which the relative humidity was about 50 percent. The results were not intended for purposes of rigorous comparison between these cords and those tested elsewhere at different times, so extensive equipment for close humidity control was not set up. In most cases, tests on individual specimens of a particular sample were staggered in such a way that observations corresponding to the same duration of creep were not made simultaneously. Thus, while moderate upward and downward drifts in the relative humidity and temperature may have had an influence on the creep behavior of individual specimens, it is believed that their effects on the measurements cancel out completely enough in the sample averages to justify the main conclusions of the present exploratory study.

The results of these measurements are shown graphically in Figs. 2 to 5. In each figure the original load, based on the cord number (or

specific linear weight) of the sample before loading, is given on the left-hand side of the figure, adjacent to the appropriate curve. In Figs. 2–4 the elongation ϵ is expressed in percentage of the original length of the unstressed cord.

Figure 2 shows the creep in a cotton cord of 17/4/3 construction under a series of loads in which the maximum is about 60 times the minimum. This cord had been given a treatment in which it was stretched in a wet condition, and then dried while still under tension. Since the creep tests extended over a period of only thirty minutes, no constant-rate flow is discernible. The behavior conforms to Eq. (1b). The graph reveals two features having general application to creep in textile materials: *viz.*, the initial deformation, ϵ_1, is clearly not proportional to the original load, or even to the initial load σ_1, nor is the value of the coefficient b, in the transient term, proportional to these quantities. In the simplified theories of creep, the quantities corresponding to ϵ_1 and b are often represented as being proportional to σ_1. Typically, Fig. 2 indicates the invalidity, from the empirical standpoint, of such generalizations. Dillon and Prettyman[15] have noted a similar lack of linearity between the slope b and the load, in their creep data on tire cords.

The results given in Fig. 3, for a regular wet-twisted 29/5/3 cord, cover a much longer period of creep (4320 minutes) than do those given in Fig. 2. The creep behavior here shows a consistent deviation from Eq. (1b). The broken lines are the graphs of equations of the type of Eq. (1b) determined by the method of least squares, while the solid lines are for the best fitting equations of the type of Eq. (1a). The particular numerical equations giving ϵ in percent are shown. The experimental points for the loads of 0.185 and 0.549 g/grex fall in better agreement with the latter curves than with the straight lines. For the remaining two loads one curve shows as good conformity with the observations as the other. All four sets of data, however, show a tendency toward greater elongations at the later time than would be predicted by the best-fitting straight lines drawn through the first eight or nine points in accordance with Eq. (1b). The constant creep rate a tends to increase with the load.

That the constant-rate component of the observed creep in Fig. 3 arises from slippage between fibers is a reasonable tentative inference. If such slippage is the cause, the constant-rate flow should be absent from creep in continuous filament cords, for here the linear coherence does not depend on inter-filament forces. To examine this point, creep observations were made on a commercial viscose rayon cord of 1100/2 continuous-filament construction. The observations, extending over nearly 13,000 minutes are summarized in Fig. 4. It will be readily seen that, though a very moderate load (710 g, or 0.290 g/grex) was used, a constant rate component is in evidence. It becomes appreciable, however,

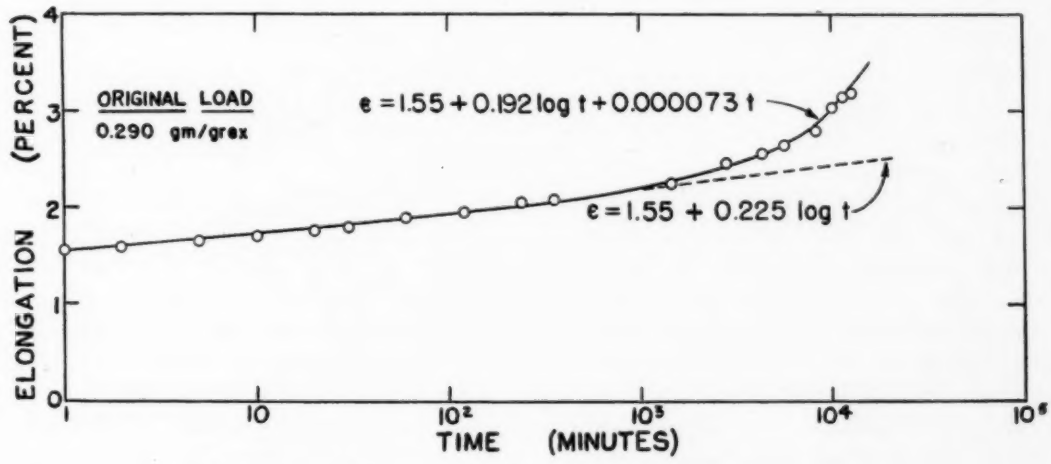


FIG. 4. Creep in a continuous-filament 1100/2 viscose rayon cord.

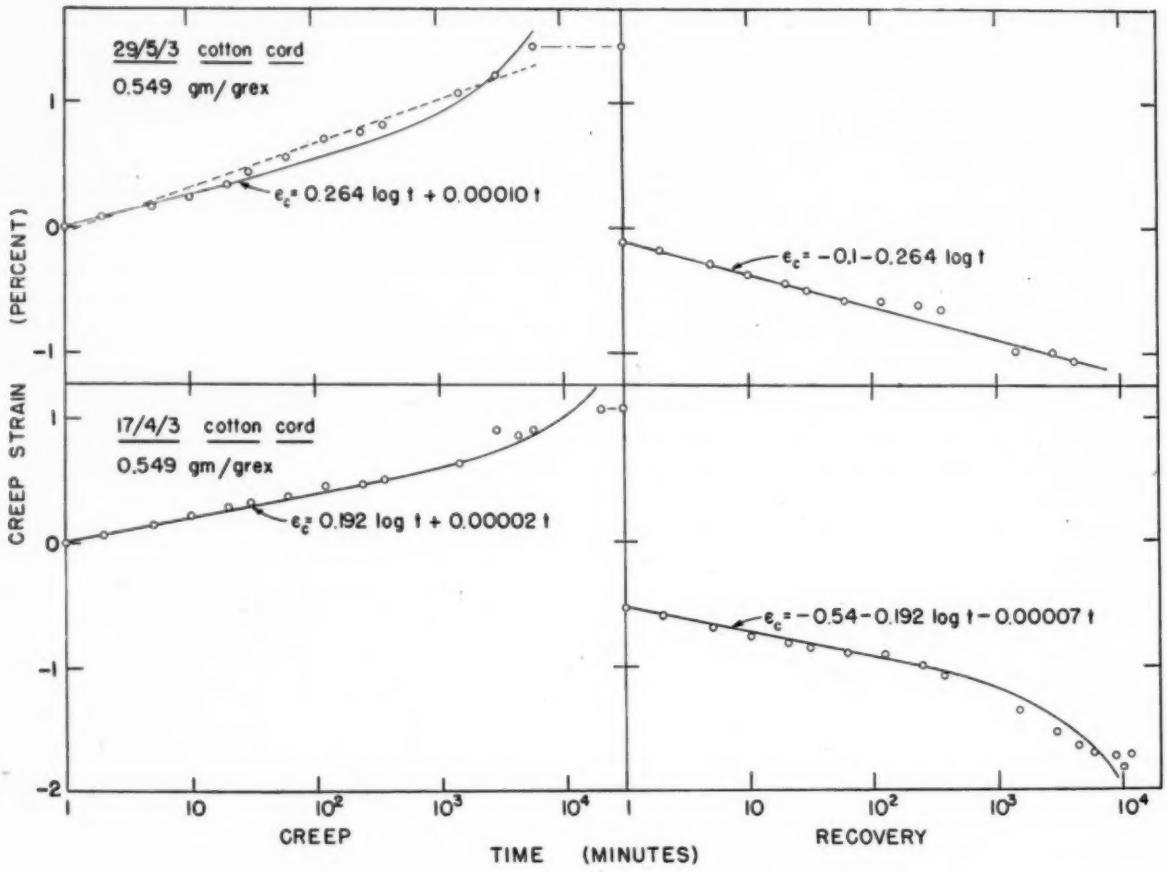

29/5/3 cotton cord
0.549 gm/grex
$\epsilon_c = 0.264 \log t + 0.00010 t$
$\epsilon_c = -0.1 - 0.264 \log t$
17/4/3 cotton cord
0.549 gm/grex
$\epsilon_c = 0.192 \log t + 0.00002 t$
$\epsilon_c = -0.54 - 0.192 \log t - 0.00007 t$
CREEP STRAIN (PERCENT)
CREEP
RECOVERY
TIME (MINUTES)

FIG. 5. Creep and recovery in a regular 29/5/3, and a dual stretched 17/4/3 cotton cord. Creep in both cords took place under an original load of 0.549 g/grex.

only after about 1000 minutes of creep. Up to about 1440 minutes (1 day), the elongation is well represented simply by Eq. (1b), as indicated by the dashed straight line. The continuous growth, which in this continuous-filament cord may be confidently taken as arising from viscous flow, proceeds at a slightly lower rate than in the cotton cord in Fig. 3, under a comparable load.

While the extensive experiments of Dillon and Prettyman[15] on cotton and rayon cords fail to show the upward curvature in the graphs of *elongation vs. logarithmic time*, the reason for this would appear to be simply the limitation on the duration of their tests. Most of their tests covered a period of about one day. As has been noted, in the present room temperature tests the constant rate flow did not become noticeable until after about 1000 minutes had elapsed; presumably such a lower time limit would apply also to elevated-temperature tests such as Dillon and Prettyman's.

In Fig. 5 are shown recovery, as well as creep data, on two cotton cords, one the 29/5/3 regular cord to which Fig. 3 refers, and the other, a 17/4/3 dual-stretched cord, treated by the process of Philipp and Conrad.[30] On completion of the creep observations, the loads were removed, except for a clamp which weighed about 30 g, and carried the index mark. The subsequent contractions of the cords were measured periodically to yield the recovery curves in Fig. 5. To condense the graphs, only the creep ϵ_c, i.e., the sum of the terms for the transient and constant rate flows, has been plotted. The points represent the means of the creep in each specimen, expressed as percentage of the particular initial

[30] H. J. Philipp and C. M. Conrad, J. App. Phys. **16**, 32 (1945).

(1-minute) length of the specimen in the creep test.

The data on the 29/5/3 cord appear to support Eqs. (1a) and (1b) equally well, for both the creep and recovery processes, if the parameters a and b are given negative signs in the application to recovery. Similarly, the creep data on the 17/4/3 cord favor neither equation, but those for the recovery in this cord show a definite conformity to Eq. (1a). The Boltzmann superposition principle[13] does not hold for these results since very little of the initial strain was recovered when the cords were unloaded. However, it is found, in partial agreement with this principle, that a single value for the coefficient of the transient term $b \log t$ can be used in the equations for both creep and recovery in both cords, to yield curves which satisfy the data moderately well. The necessary values for the rate parameter a are not found to be the same; in one case, that of recovery in the 29/5/3 cord, $a=0$, if b is given the value 0.264, which the creep data favor. A smaller value of b which, with $a=0$, would give a better fit to the recovery data, would lead to a poorer fit in the creep curve. The particular equations and the corresponding curves are shown in Fig. 5.

The appearance of a constant-rate term in the expression for recovery may be deduced from the Eyring theory. In recovery there is no constant external load producing continuous flow. However, the molecular network structure, such as exists in cellulosic fibers, provides small retractive stresses, which come into play when the external load is removed. When the stresses become very small, the rate of deformation, according to the Eyring equations, becomes sensibly constant. The tendency of recovery to approach a constant rate of flow after long intervals of time has been noted also in phenolic plastics by Telfair, Carswell, and Nason.[4]

SUMMARY

As has been shown in the foregoing discussion, there is abundant evidence for accepting Eq. (1a) as the basic, empirical law for creep under constant load. Equation (1b) is merely a special form of Eq. (1a), for the case of $a=0$. Equation (1a) has an adequate foundation in the reaction-rate theory of deformation, through which also it is related to the empirically-sound Eq. (2a) for stress relaxation. As with other general laws (Newton's viscosity law, for instance), there are materials which behave in an anomalous manner with respect to Eq. (1a), and there are limits outside which the equation does not hold. It evidently does not apply at high stresses near the breaking load, for instance. Notwithstanding their limitations, Eqs. (1a) and (2a), together with their underlying general theory, provide the framework of what may be regarded the *normal* flow behavior of amorphous solids. As the basic, empirical relations for normal elasto-viscous phenomena, Eqs. (1a) and (2a) would appear to be in the same category as Hooke's law for pure elastic and Newton's law for pure viscous phenomena.

The cited work of Busse and co-workers, Dillon and Prettyman, and Wakeham, as well as that reported in the present exploratory study, all point to the applicability of Eqs. (1a) and (2a) to the normal plastic behavior of tire cords. The equations have the usual limitations of empirical, semi-theoretical relations, in that they cannot be used validly to compute strains or stresses for periods 100 or 1000 times as long as that over which the constants were evaluated. When Eq. (1a), with the constants for tire cord, is used in such extrapolations, absurdly large elongations are obtained. It must therefore be concluded that the constant-rate flow, represented by the parameter a, does not go on indefinitely, but that after creep times on the order of 10^5 minutes the structure of the cord or its component fibers (as in the case of the acetate and Nylon samples mentioned previously) becomes sufficiently stabilized to counter-balance the residual forces producing the constant-rate flow.

In view of their wide verification, extending to non-textile materials, it would appear feasible to use Eqs. (1a) and (2a) as the basis for more or less standardized tests of creep and relaxation in cords and yarns. Just as a sample of steel is characterized by its moduli of tensile elasticity and rigidity at a certain temperature, so the plastic, or growth, properties of a cord may be identified by the values of the parameters ϵ_1, a, b, σ_1, and β under specified conditions of

load, temperature, and humidity. Dillon and Prettyman[15] have taken a step in this direction with their use of the quantities "initial compliance" and "weighted creep." There can be no doubt that the utility, for purposes of comparisons, of creep and relaxation data on tire cord and other textiles, will be broadened if they are interpreted in terms of generalizations which are valid and at the same time are generally recognized. Equations (1a) and (2a) may be adopted as such generalizations.

ACKNOWLEDGMENT

For such original observational data as are presented in this study, the author has relied on the experimental assistance of Mary L. Nelson, Hilda M. Ziifle, and Trinidad Mares, of whom the first two named also made many of the preliminary calculations.

Fatigue Failure of Rayon Tire Cord

R. C. Waller and W. E. Roseveare

Viscose Rayon Research, E. I. du Pont de Nemours and Company, Richmond, Virginia

(Received February 16, 1946)

Tire-cord fatigue tests are characterized and interpreted in terms of the following parameters: magnitude and frequency of the cyclic stroke, average load, minimum load, temperature, and humidity. By varying one test parameter at a time, it is shown that the effects of creep and degradation are minimized by using a high frequency and a high stroke. The relative rating of two samples of rayon depends on the parameters of the test and on the modulus and twist structure of the samples. Fatigue breakdown is interpreted in terms of progressive fracture under localized strains occurring at the peak stress of the cycle. A comparison of the fatigue behavior of rubbers, metals, and tire cord suggests they may all involve the same basic process of crack growth.

INTRODUCTION

UNDER repeated cyclic stresses, most solids show mechanical breakdown under stresses far below the normal breaking values. This failure of materials under cyclic deformations is called fatigue or dynamic fatigue, while the failure of a material under steady load is called creep-failure or static-fatigue.

Since 1900, many papers have appeared on the fatigue of metals. More recently, failures in plastics, elastomers, and fibers have been shown to result from continued cyclic deformations. The fatigue failure of tire cord is responsible for a type of tire failure important in certain classes of service.

The present paper discusses some aspects of the general theory of fatigue and the behavior of rayon tire cord in various laboratory fatigue tests.

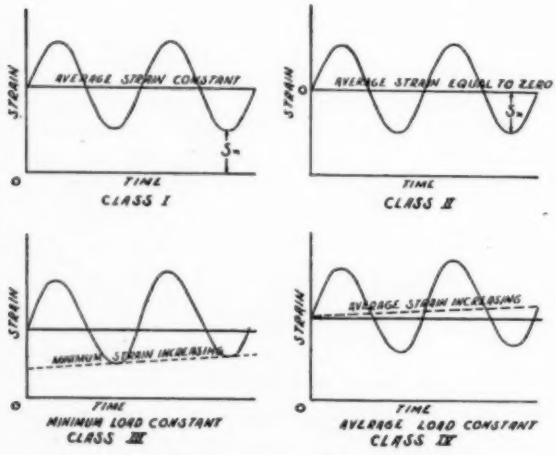


FIG. 1. Strain cycles for various fatigue tests.

FATIGUE TESTS

The rate of fatigue is affected by the structure of the sample, presence of chemically reactive materials, humidity, temperature, period, and magnitude of the localized cyclic strain, and the local magnitude of the minimum strain at the end of a cyclic stroke.

Figure 1 shows the strain throughout the cycle

for four classes of constant-stroke fatigue tests. These classes differ in the ratio of cyclic strain to average strain and in the changes of average strain from one cycle to the next. These classes of fatigue tests are different from the tests with constant cyclic stress, as usually used for metals. The Class I test of Fig. 1 involves a constant cyclic elongation with a constant minimum strain, S_m, throughout the fatigue life. The Class II test, having equal compression and elongation, is a special case of Class I in which S_m is negative and equal to one-half the cyclic displacement.

Instead of keeping the minimum strain constant as in Classes I and II, one may adopt a Class III test by taking up the slack in a sample under a constant minimum load by means of a ratchet mechanism operating at the point of minimum strain in the cycle. The cyclic deformation remains constant, while the total length of the sample increases, as shown by the increase in minimum strain for Class III of Fig. 1.

Another way of taking up the slack is to hang a weight on the sample, while vibrating the other end up and down with the weight remaining practically stationary. This procedure maintains a constant average load, which likewise causes the sample to grow in length, as shown for Class IV of Fig. 1. The essential difference between the tests of Classes III and IV is in the minimum load being constant in Class III, while the maximum load is constant in Class IV; however, the samples in both tests grow in length until failure.

As either the minimum load in Class III or average load in Class IV is increased, these tests approach a static-fatigue or creep-failure test. As the stroke is increased, these tests become more sensitive to elongation characteristics or elastic limit of the material. A stroke may be chosen that will break a low elongation material in the first stroke and yet not break a high elongation material in many thousand cycles; under such conditions the test will cease to be a fatigue test for the lower elongation material. Therefore, the use to be made of the results of the test must be considered in selecting the conditions for a fatigue test.

A fatigue test may have two purposes, namely, an over-all fatigue rating under service conditions, or a measurement characterizing a sample in terms of what we may call a "true fatigue" resistance which is independent of creep under static load, heat-aging, and deteriorating agents. This true fatigue will be a function of the temperature, moisture content, stroke, and structure of the sample. A complete fatigue rating involves both a rating under standard conditions and the sensitivity of this rating to each of the conditions of which the fatigue test is a function. To obtain practical relative fatigue ratings of materials having widely different properties, such as glass, cotton, rayon, and Nylon, it is necessary to duplicate closely the degree of strain occurring in service failure. Since service conditions are varied and difficult to define, one chooses what appears to be a reasonable set of parameters for a laboratory fatigue test. For narrow classes of materials having similar elongations and temperature sensitivities, etc., such measurements should correlate with service fatigue.

Class I and II tests have been used on rubber[1]

TABLE I. Fatigue tests.

Test	Class	Stroke % elong.	Load in grams per denier	Temperature	% regain moisture in cord	Frequency in c.p.m.
A	III	1.6	0.14 (Min.)	49°C	3.7	360
B	IV	1.6	1.0 (Av.)	100°C	2.0	3000
C	IV	0.5	0.65 (Av.)	170°C	Dry	3650

TABLE II. Properties of yarns at 23.9°C (75°F) (3″*Z*″ twist wet or 60 percent relative humidity).

Cord No.	Tenacity g.p.d. Cond.	Tenacity g.p.d. Wet	% elongation Cond.	% elongation Wet
1	3.62	2.25	10.1	20.3
2	3.70	2.30	9.8	19.6
3	3.56	2.12	9.8	17.9
4	5.23	3.81	7.4	7.8

TABLE III. Properties of cords at 23.9°C) (75°F) (1100/2 construction 15″*Z*″–11″*S*″ twist).

		Conditioned at 60% relative humidity			Oven-dry cord		
Cord No.	Denier	Strength in lb.	% elong. at 10# load	% elong. at break	Strength in lb.	% elong. at 10# load	% elong. at break
1	2492	16.4	9.0	16.5	20.1	4.1	12.1
2	2412	16.0	8.1	13.5	19.0	3.9	10.5
3	2449	15.1	8.1	13.1	17.8	4.0	9.8
4	2439	18.6	5.0	9.4	21.7	2.1	6.4

[1] S. M. Cadwell, R. A. Merrill, C. M. Slowman, and F. L. Yost, Ind. Eng. Chem. Anal. Ed. **12**, 19 (1940).

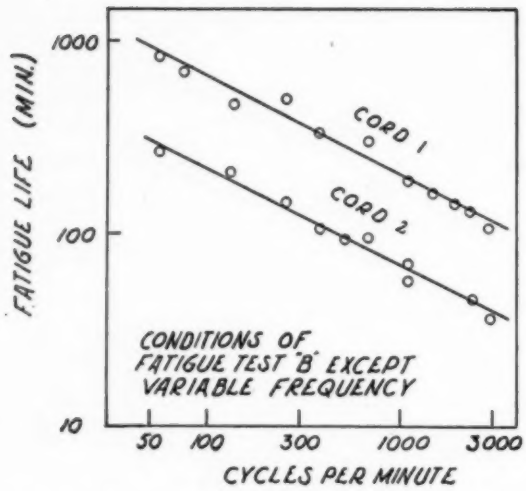


FIG. 2. Effect of frequency of stresses on fatigue life.

and a Class II test has been designed by Ray[2] for tire cord embedded in rubber. This latter test gives an over-all fatigue rating for the service conditions simulated by the test; however, the results of this test depend on the technique and materials used in embedding the cord in rubber, as well as on the inherent fatigue resistance of the cord materials. Class III and IV fatigue tests have been developed for tire cord by Castricum[3] and Busse and others,[4] respectively. Class I and II tests are much more difficult to apply to tire cord free of rubber than the Class III and IV tests. To simplify both the technique and the interpretation of the results, the present investigation has been confined to the Class III and IV tests.

Experimental Tests

The conditions of some of the selected standard tests are given in Table I. The load is in grams per denier, and the frequency is in cycles per minute.

The effect on fatigue life of varying each of these parameters except stroke, one at a time, has been studied for one or more of these fatigue tests in order to find conditions where the fatigue life is independent of heat-aging and least dependent upon creep under static load. The properties of the cords used in this study are given in Tables II and III.

All tests were made on a Scott IP4 tester using a loading rate of 16 g.p.d. per minute. Cords 1, 2, and 3 represent three different types of commercial viscose rayon tire yarns, while Cord 4 represents a low elongation viscose rayon yarn.

The coefficient of variation for 54 determinations in Test *A* was 20 percent. For a normal distribution and a probability of 90 percent, this corresponds to an uncertainty of 18 percent for six observations and 11 percent for 12 observations. Most values reported in this paper represent averages of either six or 12 observed values.

EXPERIMENTAL RESULTS

Effect of Frequency of Stresses

The effect of the frequency of cyclic stresses on the fatigue life (time to failure) of Cords 1 and 2 may be seen in Fig. 2, where frequency and fatigue life are plotted on logarithmic coordinates. A straight line has been drawn through the points from 60 to 3,000 c.p.m. The relationship represented by the straight line may be written in the following form

$$F = k\omega^{-m}, \tag{1}$$

where F is fatigue life in minutes, ω is frequency, and k and m are constants. The fatigue life is proportional to the frequency raised to the negative fractional power $-m$; the number of cycles to failure is equal to the product of the life and the frequency and, therefore, proportional to the frequency raised to the positive fractional power $1-m$. Busse and others[4] re-

TABLE IV. Irreversibility of fatigue processes in fatigue test *A*.

Initial fatigue time (hours)	Intermediate treatment	Fatigue time after treatment (hours)	Total fatigue life (hours)
4.7	Control: fatigued direct to failure	3.6	8.3
4.7	Conditioned 24 hours at 60% R.H. 23.9°C	4.2	8.9
4.7	Wet out, dried on jig at original length	3.4	8.1
4.7	Heated for 1.5 hours at 150°C in a dry evac. tube	3.8	8.5

[2] F. Ray, U. S. Patent No. 2,235,622.
[3] M. Castricum, U. S. Patent No. 1,923,296.
[4] W. F. Busse, E. T. Lessig, D. L. Loughborough, and L. Larrick, J. App. Phys. **13**, 715 (1942).

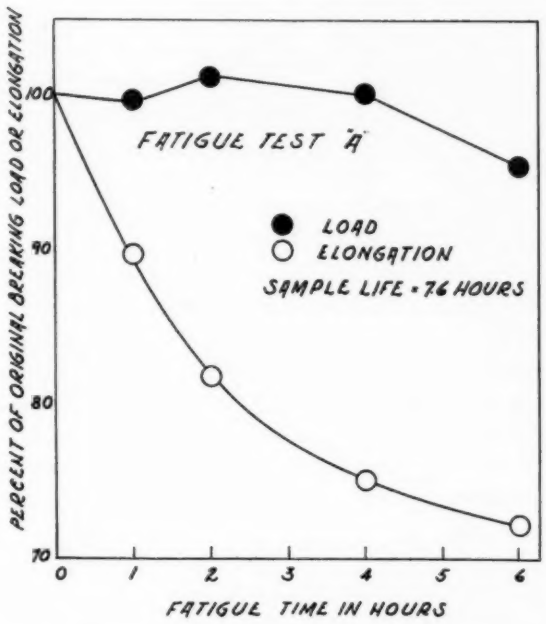


FIG. 3. Effect of fatiguing on breaking load and elongation of a tire cord.

ported the fatigue life of rayon and cotton cords to be independent of the frequency; this was due probably to the heat-aging involved in their test and to the limited frequency range wherein the stress-strain behavior of their tester was unaffected by resonance. No dispersion or critical frequency ranges have been observed, indicating that no critical relaxation frequencies of the cellulose occur in the frequency range employed in this study. The values of m in Eq. (1) are the same for Cords 1 and 2 within the limits of experimental error. Increase in the fatigue life with increasing frequency was reported for rubber by Yost,[5] although the functional relationship was not given.

Irreversibility of Fatigue Processes

The fatigue processes involved in Test *A* appear to be irreversible. Samples of Cord 3 having an average fatigue life of 8.3 hours were fatigued for approximately one-half of this time, given an opportunity to recover, and then fatigued to failure. The total time of fatiguing and the kind of intermediate recovery treatment used are given in Table IV. None of the fatigue resistance expended in the initial period was recovered by the various treatments tried; the differences in total fatigue life are within the experimental error. Busse and others[4] have shown that the processes of Test *C* are not reversed by treatments such as mechanical flexing, conditioning, or wetting with various solvents.

The effect of fatiguing, for various times short of failure, on the breaking load and breaking elongation of a tire cord having a fatigue life of 7.6 hours is shown in Fig. 3 for Test *A*. There is no appreciable change in the breaking load until near the end of the fatigue life; however, the breaking elongation rapidly decreases during the first part of the fatigue life.

Growth during Fatigue Testing

The repeated deformation of a cord results in the production of a permanent elongation. A continuous record of elongation *vs.* time is given in the upper, left part of Fig. 4 for a sample of commercial tire cord in Test *A*. The cords grow very rapidly in the first few minutes, after which the rate of elongation becomes proportional to the logarithm of the time as shown by the straight line part of the curve in the lower right part of Fig. 4. The stage just before failure is characterized by a period of rapid growth. Similar straight line plots are obtained for a large part of the creep of a cord under static load. Leaderman[6] has observed similar behavior

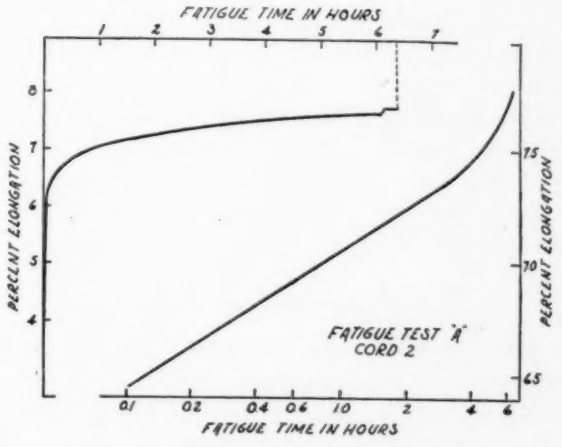


FIG. 4. Growth behavior during fatiguing.

[5] F. L. Yost, Trans. A.S.M.E. **65**, 881 (1943).

[6] H. Leaderman, *Elastic and Creep Properties of Filamentous Materials and other High Polymers* (The Textile Foundation, 1943).

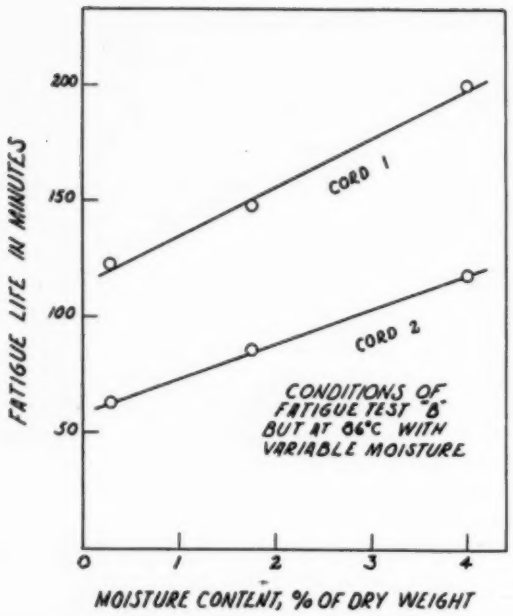


FIG. 5. Effect of moisture on fatigue life.

of the creep under static load for various textile fibers while Chasman[7] has obtained such creep curves for plastic sheets.

Effect of Moisture on Fatigue

Under conditions approaching those in fatigue Test *B* where heat degradation is insignificant the fatigue life is increased by increasing the moisture content of the cord. Figure 5 shows a plot of the fatigue life of two types of tire cord as a function of the moisture content of the cord obtained on desorption. This effect of moisture is the opposite of the effect of moisture on the life of a cord under a static load. The increased fatigue lives at the higher moisture contents probably result from the lower cyclic stresses present in moist cord. Increasing the moisture content of cord in a tire is undesirable, since it increases the rate of growth of the tire.

Effect of Temperature on Fatigue

The effect of temperature on life has been determined with conditions other than temperature as in Test *B* and Test *C* and for creep failure under two different static loads. It was impractical to surround the cords with completely dry air or to provide a constant moisture content in the cord at each temperature. The air supplied to the testers for these temperature coefficient studies contained 60 percent relative humidity at 23.9°C with other conditions as in Test *C* or 13 percent relative humidity at 23.9°C with other conditions as in Test *B*. As the temperature of the test is lowered, the moisture content of the cord will increase. Only below 80°C will sufficient moisture be present to affect the life with other conditions as in Test *B*.

The results of these temperature-coefficient measurements are given in Fig. 6. The lives in both the creep and fatigue tests are exponential functions of the reciprocal of the absolute temperature. For each fatigue test, the temperature coefficient is nearly the same for the two types of high-tenacity rayon yarns.

It follows from the Arrhenius equation giving the activation energy of a physical or chemical molecular rate process[8] that

$$\log_e L = (E/RT) + C, \qquad (2)$$

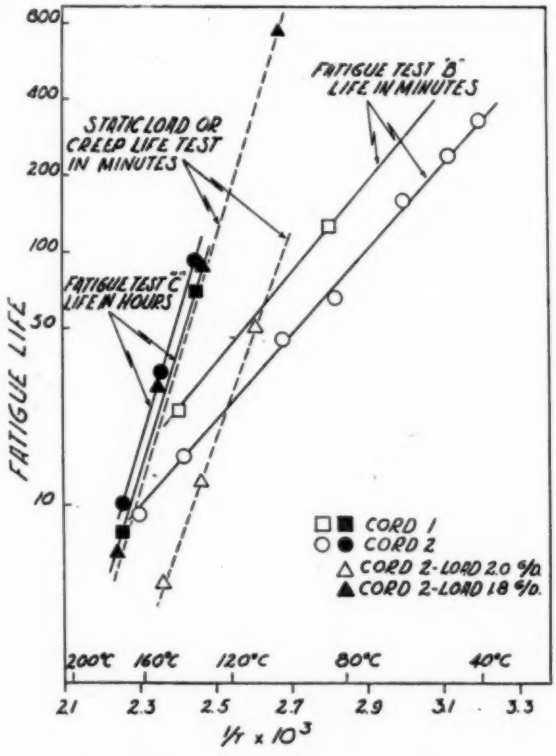


FIG. 6. Effect of temperature on fatigue life.

[7] B. Chasman, Modern Plastics **21**, 6, 145 (1944).

[8] S. Glasstone, K. J. Laidler, and H. Eyring, *Theory of Rate Processes* (McGraw-Hill Book Company, Inc., New York, 1941).

where L is the time required for a selected amount of change to take place, E is the activation energy for the process, and C is a constant. Table V gives the activation energies determined from the slopes of the straight lines in Fig. 6. The value of the activation energy reported in the table for the fatigue test of Busse and others[4] is a corrected value calculated from the original data.[9] This test approaches the conditions of Test C, and the activation energies of these two tests are in reasonable agreement.

Table V also gives a rough activation energy determined in our laboratory[10] for the degradation of Cord 2 in an oven containing air. This activation energy was calculated using the times for equal loss in strength on heating rayon cord at 160°C and 185°C in an oven. The difference between the activation energies for the two creep tests is within experimental error. The activation energy for Test C agrees, within experimental error, with values for creep and with the value for the thermal degradation of tire cord under conditions similar to those in Test C. On the other hand, the comparatively low activation energy found for fatigue Test B indicates that the most important process occurring in this test is not the same as that involved in either heat-aging or creep under static load. In addition, the large difference between the activation energies for Tests B and C suggests that their rates are determined by basically different processes.

The conditions for Test C are more likely to involve heat-aging than the conditions of Test B, since the time available for aging at the high temperature is larger by 600 percent. The importance of heat-aging and creep in Test C cannot be reduced by merely lowering the temperature, since the temperature coefficient for this test is the same as that for heat-aging and for creep. On the other hand, the effects of heat-aging and the creep under static load for Test B become relatively less important as the temperature is lowered. Since straight lines are obtained for the Test B conditions of Fig. 6, it appears that neither of these factors is important for this test even at the highest temperature studied.

TABLE V. Activation energies.

Cord test	Activation energy in calories/kinetic unit
Fatigue Test B	7800
Fatigue Test C	23000
Fatigue Test of Busse *et al.*	19000
Creep (1.8 g.p.d.)	22000
Creep (2.0 g.p.d.)	19000
Thermal degradation of tire cord	25000

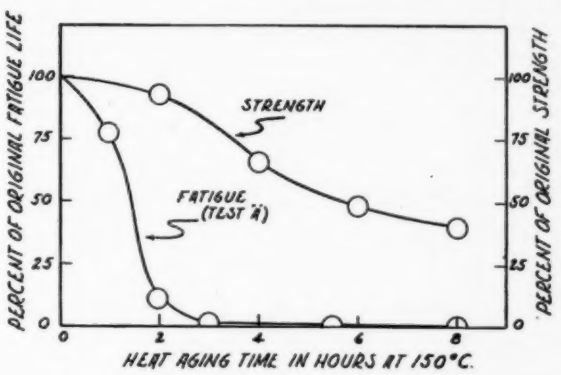


FIG. 7. Effect of heat degradation on fatigue life.

Considerable heat degradation is in evidence in cords fatigued in air under the Test C conditions. This is indicated by discoloration of the test samples. Under the conditions of temperature, moisture, and time in the Type C fatigue test, an unloaded cord may lose 25 percent of its strength because of heat degradation in a normal failure time. Fatigue life is particularly sensitive to previous heat degradation, as shown by the data in Fig. 7. Two hours' heating at 150°C, in sealed tubes containing air and 0.12 gram of moisture per gram of cellulose, reduces the fatigue life by 90 percent, while the loss in tenacity is relatively small. This parallels the behavior of metals where corrosion may greatly reduce fatigue life without greatly changing the tensile strength.

Effect of Load on Fatigue

The life under the conditions of fatigue Tests B and C except for variable load is an exponential function of the average load on the cord. The logarithm of fatigue life is plotted against the average load in Fig. 8. The plot shows that the two samples have the same fatigue life in a Type

[9] W. F. Busse, E. T. Lessig, D. L. Loughborough, and L. Larrick, J. App. Phys. **16**, 120 (1945).
[10] E. A. Tippetts and A. W. Powell, unpublished data.

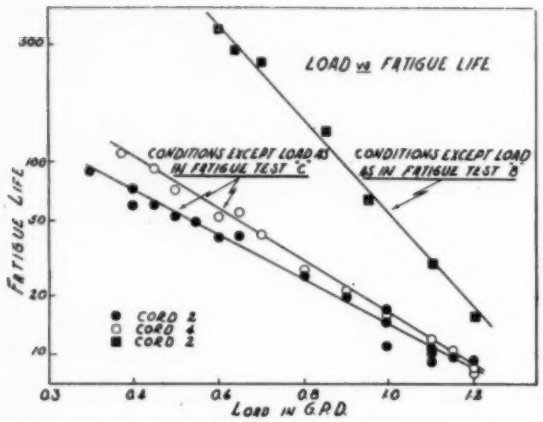


FIG. 8. Effect of load on fatigue life.

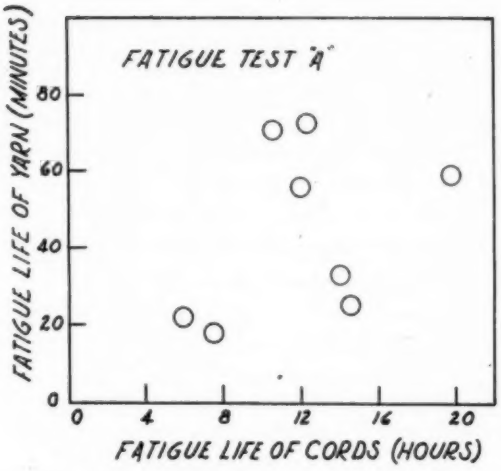


FIG. 9. Fatigue life of yarns *vs.* cord.

C Test at an average load of 1.2 g.p.d.; however, the fatigue resistance of Cord 4 is superior to that of Cord 2 at low loads. A change in the relative fatigue ratings of cotton and rayon on changing the load has been reported by Busse and others[4] for a test similar to fatigue Test *C*. Cords tested at lower average loads are more affected by heat-aging than are those tested at higher loads. This effect may be responsible for the change in relative fatigue ratings observed. Similar transpositions of fatigue ratings at high and low loads are well known in the case of metals. It has been shown for metals[11] and cellulose acetate sheets[12] that cyclic stresses below the "endurance limit" of these materials do not produce failure in an infinite number of cycles. At stresses above the endurance limit, the fatigue life is proportional to the applied stress raised to a fractional power. It is possible that if the effects of creep under static load and heat-aging were eliminated, the fatigue life of rayon might follow a similar relationship.

Effect of Stroke and Elongation Characteristics of the Sample

A small change in the stroke makes a great difference in the fatigue life. Changing the stroke of Test *A* from 1.6 percent to 1.7 percent decreased the fatigue life of a commercial tire cord by 40 percent. Decreasing the modulus of a cord has an effect on the fatigue life similar to decreasing the stroke. The modulus of an 1100/2 rayon cord is usually reported in terms of its reciprocal, the elongation of the cord under 10-pound load. A yarn dried under conditions permitting shrinkage gives a cord with a high elongation under a 10-pound load when compared with a cord obtained from a yarn dried under tension. Such an increase in fatigue life may be considered, in part, artificial, since increasing the 10-pound load elongation above the normal value for a commercial tire-cord yarn is usually accompanied by an objectionable increase in the initial rate of enlargement or growth of the tire under the inflation pressure. For this reason, fatigue ratings of various productions are usually made using cords of equal 10-pound load elongations. If the elongation under 10-pound load deviates a small amount from the specified value, an empirical correction factor may be applied to correct the fatigue to the standard elongation.

Cords with inherently different elongation characteristics will have relative fatigue ratings which depend on the stroke and other conditions of the fatigue tests as Table VI indicates. Cord

TABLE VI. Comparison of ratings on Tests *A* and *C*.

Sample	Conditioned elong. at 10# load (see Tables II and III)	Fatigue rating Test *A*	Fatigue rating Test *C*
Cord 2	8.1	190	90
Cord 3	8.1	100	100
Cord 4	5.0	3	120

[11] D. Landau, *Fatigue of Metals* (The Nitralloy Corporation, 230 Park Avenue, New York, New York).
[12] W. N. Findley, Modern Plastics **18**, 9, 57 (1941).

4, a high-strength, low elongation viscose-rayon cord, compares unfavorably with commercial tire Cords 2 and 3 in Type *A* Test, while it is superior in a Type *C* Test. This appears to be due to the lower stroke of Test *C*, which measures largely creep and heat resistance.

Fatigue of Filaments, Yarns, and Cords

The fatigue life of a single filament measures its resistance to cyclic stresses along the fiber axis, but when a number of filaments are incorporated into a twisted structure such as a cord, additional possibilities for failure are introduced by the interaction between the filaments. Since tire-cord rayons are highly anisotropic, a fatigue failure resulting from the transverse forces present in a twisted structure may be of a different character than one produced by cyclic stresses along the fiber axis. Single filaments and zero-twist yarn were tested, but the large spread in these results necessitated the use of a low twist yarn (3*Z* twist) to measure the longitudinal fatigue resistance.

The fatigue life of yarn (3*Z* twist) is plotted against that of cord (1100/2) in Fig. 9 for the fatigue Test *A*. This test shows a longer life for the cord than for the yarn, while in the fatigue Test *C* the yarn has the greater life. The longitudinal fatigue resistance, as measured on a low twist yarn, does not appear to correlate with the fatigue resistance of its cord. This suggests that the relative fatigue ratings of tire yarns may be a function of cord construction. It is well known that the absolute fatigue life of cords from a given type of yarn is also a function of cord construction. The optimum cord construction for one type of rayon yarn is not necessarily the best construction for another type of rayon yarn having different properties.

Effect of Parameters of Test on Fatigue Rating

The ratings of a series of tire cords determined on Test *B* correlate very well with those from Test *A*, as shown in Fig. 10. These two tests have strokes with the same percent elongation, but they differ considerably in temperature, frequency, and moisture content of the cord. Figure 10 shows that the results of Test *C*, using a low stroke and very high temperature, correlate very poorly with the results of the other two fatigue tests. The test parameters to be used in determining a fatigue rating should approach service conditions as closely as possible without making the test life too long. For rayon tire cord, the conditions of the Type *B* Test give a reasonably short fatigue life in which creep under static load and heat-aging are relatively unimportant. A test combining fatigue, heat-aging, and creep under static load would give more nearly a service rating for a cord if one could make these factors of the same relative importance in the test as in the particular type of tire service for which the tire cord is being evaluated.

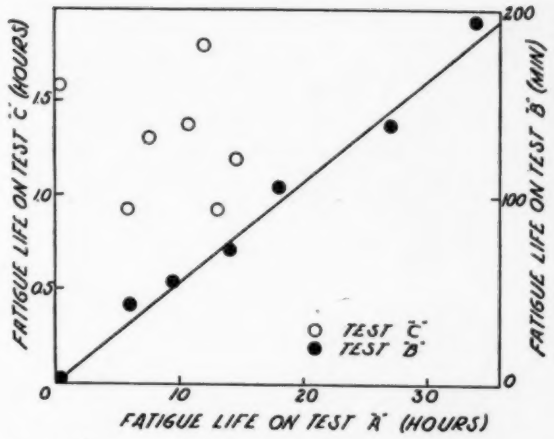


FIG. 10. Relative fatigue ratings with three different tests.

DISCUSSION OF THE MECHANISM OF FATIGUE FAILURES

A rayon cord fatigued to incipient failure shows a large number of broken filaments all along the cord. Such a cord shows little loss in strength. Relatively few broken filaments are observed along the length of a cord broken in a tensile tester. These observations suggest that the fatigue occurs in varying degrees at many points in the individual filaments making up the cord. Fatigue failure of the cord results after a critical number of filaments have failed near a point in the cord so that the remaining filaments rupture rapidly because of increased load per filament at this weak spot.

In a cord or yarn it is impossible to watch crack growth in the individual filaments. Some partially broken filaments have been observed in samples of fatigued cords. Fibrillation of the

FIG. 11. Photomicrographs of filaments from fatigued cords.

filaments has been observed in fatigued cords in unpublished observations of L. M. Welch in this laboratory. Some of his photomicrographs of filaments from fatigued cords made from typical tire yarns are shown in Fig. 11. Many filaments throughout the fatigued cord show surface fibrillation. Examination of cords removed from the point of maximum fatigue of a tire showing fatigue failure has failed to show this surface fibrillation. This may be due to the difficulty in removing filaments from the partially impregnated cord of a tire and to reduction in the slippage of one filament over another by the adhesive coating which cements the surface filaments together. Cords broken in a tensile tester do not show surface fibrillation and the broken ends of the filaments show less longitudinal fracturing than the ends of filaments from fatigued cords. These observations suggest that fatigue of tire cord is associated with an internal fracturing of the filaments and that surface fibrillation is only important in cords not coated with adhesive.

The initiation of cracks should be more frequent for the higher peak stresses. Increasing the stroke, the average load, or the modulus of the cord reduces the fatigue life. The effect of the latter factor is true whether the modulus is increased by increasing the frequency of the cyclic stroke, decreasing the moisture content of the cord, or decreasing the 10-pound load elongation of the cord. The use of a fatigue test with a constant peak load, as one uses usually for metals, should reduce the effect of the above variations in modulus on fatigue life.

The temperature coefficient of fatigue is practically zero for most metals. Deviations from this rule have been explained as due to the test sample becoming a different material with different strength properties on changing the temperature.[13] For normal metals, there appears to be practically no variation of the critical rupture stress with temperature; zinc and bismuth show almost no change in this property from liquid hydrogen temperatures up to room temperatures.[14] A comparison of the behavior of tire cord with metals raises the question as to whether the change in fatigue life with temperature may not be due to a change in the nature of the fiber. The strength of nearly dry rayon cord decreases as the temperature is raised.[15] If fatigue life is as sensitive to strength as it is to average load in Test *B*, a large part of the change in fatigue life with temperature can be accounted for as being due to change in strength. In addition, changes in the modulus with temperature should affect fatigue life. The available information does not exclude the possibility that the basic process in the fatigue of tire cord may have a zero temperature coefficient as one finds for metals. If the temperature coefficient of the rate determining step is zero for tire cord, then this step must involve the simultaneous rupture of a large number of bonds between aggregates of molecules. These bonds may include one or more of the following types: van der Waals, hydrogen, or chemical bonds. Crack growth by such a mechanism appears to be the most satisfactory explanation covering the fatigue of all types of materials.

The origin of a crack will be the point of maximum microscopic strain. The crack will tend to grow in the direction of the larger adjacent strains. A crack may run into a region of low strain whereupon this crack may cease to grow while new cracks are being formed else-

[13] M. L. E. Oliphant, Proc. Roy. Soc. **171**, 79 (1939).
[14] F. Seitz, *The Physics of Metals* (McGraw-Hill Book Company, Inc., New York, 1943), p. 164.
[15] J. H. Dillon and I. B. Prettyman, J. App. Phys. **16**, 159 (1945).

where. A large perfect crystal or even a region of greater order would be a region of low strain, and therefore the presence of such material may retard crack growth. This picture of fatigue predicts the very large observed effect of minute flaws or "stress risers" on fatigue life; it is well known that a high polish on metals greatly increases their fatigue resistance.

In the case of rubber, neoprene, and Butyl B, Cadwell and others[2] showed that the fatigue life increases to a maximum and then decreases as the minimum elongation of the test is increased. After reaching the maximum, the fatigue life decreases because the effect of increased load is more important than increased crystallization. Fielding[16] has shown that no such increase in fatigue life is found for either Buna-N or GR-S synthetic elastomers. Since this latter group does not show a fiber type of x-ray pattern on stretching, while the first group does, it is probable that crystallization is responsible for the improvement in the fatigue of the first group and that crystals retard crack growth.

Isotropic cellulose and rubber have been shown by Hermans[17] to possess many similarities in their mechanism of deformation. Rayon yarns, on the other hand, are oriented and quite crystalline so that a maximum fatigue life on stretching yarns would not be expected as in the case of rubber. High tenacity viscose rayon may be considered somewhat like stretched rubber, and, like rubber, it will be in the region where increased stretch would rapidly decrease the observed fatigue life by reducing the elastic limit of the substance. Material with a low elastic limit may have excellent fatigue under conditions involving low stroke. A highly oriented rayon may be like rubber at 500 percent minimum stretch, where rubber has an excellent fatigue life, but on doubling the stroke the fatigue life of rubber is decreased by a factor of over 100. A less crystalline rayon may correspond to rubber stretched 300 percent, where doubling the stroke decreases the fatigue only threefold.

In addition to the gross structural factors affecting fatigue, there are rate factors which affect the behavior of a material in a fatigue test. If a material is very hard, like glass, at normal temperature so that strains are slowly relieved, it will rupture quickly once the elastic limit is exceeded. A very ductile material may be deformed above its elastic limit without fracturing, since the strains are rapidly relieved. Such a material on repeated cyclic stressing under load may be very highly deformed without producing a fatigue rupture, and it is characterized by showing a large creep under high static loads. In some tests it may have a very high observed fatigue life. Rayon is intermediate between very hard and ductile materials.

CONCLUSION

The type of degradation occurring in a fatigue test is determined by the parameters of the test, and varying the values of these parameters may change the order of rating for some samples. Parameters of a fatigue test for tire-cord rayon have been chosen to minimize the effect of heat degradation and creep under a static load.

Metals, plastics, elastomers, and fibers have in their fatigue behavior similarities suggesting progressive fracture as a common basic mechanism for all fatigue failures.

ACKNOWLEDGMENT

The authors wish to acknowledge the valuable assistance of R. E. Spain in obtaining the experimental data for this paper.

[16] J. H. Fielding, Ind. Eng. Chem. **35**, 1259 (1943).
[17] P. H. Hermans, J. Phys. Chem. **45**, 827 (1941).

Dynamic Characteristics of Rubber Supports from Vibration Table Data

ERNEST G. CHILTON*
Firestone Industrial Products Company, Akron, Ohio
(Received February 11, 1946)

Measurements of the vibration of a free mass which is isolated by a rubber support from a surface vibrating at constant amplitude permit calculations of the dynamic characteristics of the support. This method permits high loadings and the application of considerable force.

INTRODUCTION

THE determination of dynamic characteristics of rubber, both natural and synthetic, has been described by a great number of authors. Values of dynamic rate (or modulus) and internal friction (or damping) in both compression and shear have been determined under many conditions.[1-3]

However, the usual laboratory apparatus used to find these characteristics in the raw material is rarely capable of withstanding the necessary dead loads, or applying forces of the order of those experienced in service. Gehman[4] and Kosten[5] have described methods based on the centrifugal force of revolving eccentric masses.

The method described here utilizes a "vibration table," a horizontal surface made to vibrate vertically with a certain pre-set amplitude and continually variable frequency. Upon this surface is mounted the specimen under test, and it, in turn, supports the desired load. By recording the amplitude of vibration of this load it is easy to determine the transmissibility of the specimen at the particular frequency. Resonance, however, the condition of maximum transmissibility, is usually chosen as the point at which physical characteristics are determined, for at this point calculations become fairly simple. It will be shown in the derivation how both dynamic rate and internal friction can be determined from the values of resonant frequency and load amplitude together with the known constants of the system.

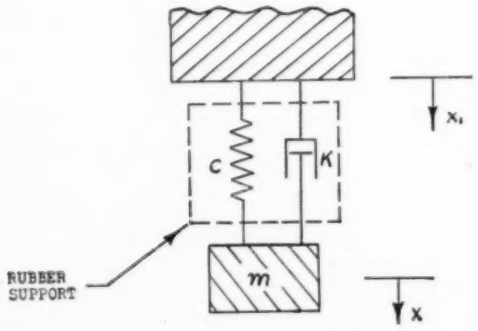


FIG. 1. Schematic diagram of vibration system.

DERIVATION OF EQUATIONS

Although the motion of the table surface is not perfectly simple harmonic, it is close enough to assume it to be of that form.

Let, therefore, the table motion be described by the equation:

$$x_1 = a \cos \omega t,$$

where x_1 = the distance of the table surface from its neutral position at time "t," a = the amplitude, or maximum value of "x_1," and ω = the frequency of the motion.

The force due to the inertia of the load is given by the product of the mass m of that load and its acceleration d^2x/dt^2, where x is the displacement of the load from its neutral position at any time (see Fig. 1).

The force due to the dynamic rate of the specimen is proportional to the relative displacement of its ends, and thus to the difference $x - x_1$. While the internal friction force is similarly proportional to the difference in velocities,[6]

$$dx/dt - dx_1/dt.$$

* Now on leave of absence at Stanford University, California.

[1] S. D. Gehman, D. E. Woodford, and R. B. Stambaugh, Ind. Eng. Chem. **33**, 1032 (1941); Rubber Chem. Tech. **14**, 842 (1941).

[2] R. B. Stambaugh, Ind. Eng. Chem. **34**, 1358 (1942); Rubber Chem. Tech. **16**, 400 (1943).

[3] J. H. Dillon, I. B. Prettyman, and G. L. Hall, J. App. Phys. **15**, 309 (1944).

[4] S. D. Gehman, J. App. Phys. **13**, 402 (1942); Rubber Chem. Tech. **15**, 860 (1942).

[5] C. W. Kosten, Proc. Rubber Tech. Conf., London 987 (1938); Rubber Chem. Tech. **13**, 381 (1939).

[6] Viscous friction has been assumed in these calculations for reasons of simple mathematical treatment as in references 1-4.

Hence, with due regard for direction, the equation of forces of the system can be written as:

$$m\frac{d^2x}{dt^2}+k\left(\frac{dx}{dt}-\frac{dx_1}{dt}\right)+c(x-x_1)=0, \qquad (1)$$

where k is the damping or friction constant and c is the dynamic rate or modulus. Substituting for

$$x_1=a\cos\omega t$$

and

$$dx_1/dt=-a\omega\sin\omega t,$$

we get:

$$md^2x/dt^2+kdx/dt+cx=ca\cos\omega t-ka\omega\sin\omega t. \qquad (2)$$

The transient part of the solution of this differential equation is of little interest as it describes the first few moments of motion before the steady state has been reached. The steady-state solution, however, is of the form:

$$x=M\sin\omega t+N\cos\omega t. \qquad (3)$$

Differentiating this:

$$dx/dt=M\omega\cos\omega t-N\omega\sin\omega t,$$
$$d^2x/dt^2=-M\omega^2\sin\omega t-N\omega^2\cos\omega t=-\omega^2x.$$

Substituting in (2) and letting ωt assume values of 0 and $\pi/2$, respectively, we get:

$$(c-m\omega^2)M-k\omega N=-ka\omega,$$
$$(c-m\omega^2)N+k\omega M=ca.$$

Using $n=k/2m$ and $p^2=c/m$ this becomes:[7]

$$(p^2-\omega^2)M-2n\omega N=-2an\omega,$$
$$(p^2-\omega^2)N+2n\omega M=p^2a.$$

Solving these equations for M and N:

$$M=\frac{2na\omega^3}{(p^2-\omega^2)^2+4n^2\omega^2}, \qquad (4)$$

$$N=\frac{a(4n^2\omega^2+p^4-p^2\omega^2)}{(p^2-\omega^2)^2+4n^2\omega^2}. \qquad (5)$$

By setting $dx/dt=0$ it is easily shown that the maximum value of x:

$$x_{\max}=(M^2+N^2)^{\frac{1}{2}}=\text{amplitude}=A.$$

Substituting for M and N from Eqs. (4) and (5):

$$A=\frac{a(4n^2\omega^6+16n^4\omega^4+p^8-2p^6\omega^2+p^4\omega^4+8n^2\omega^2p^4-8n^2\omega^4p^2)^{\frac{1}{2}}}{(p^2-\omega^2)^2+4n^2\omega^2}. \qquad (6)$$

At resonance, when A is a maximum,

$$p=\omega^8 \qquad (7)$$

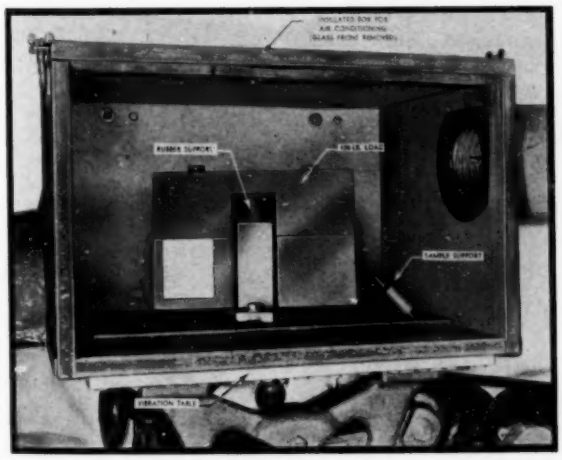


FIG. 2. Typical vibration support with 100-lb. load on vibration table.

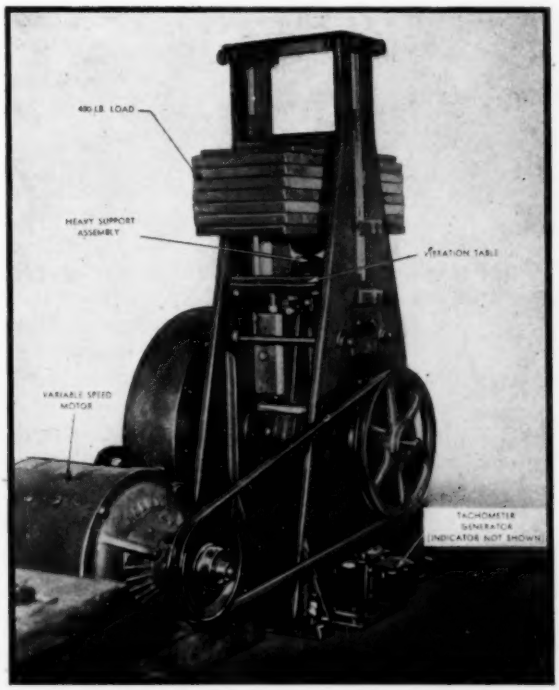


FIG. 3. Vibration support with 400-lb. load in heavy-duty vibration machine.

[7] S. P. Timoshenko, *Vibration Problems in Engineering* (D. Van Nostrand Company, Inc., New York, 1928), pp. 23 and 26.

[8] This is strictly true only for $n\to 0$; it can, however, be shown by setting $dA/d\omega=0$ that $p=\omega$ is in error only by a term proportional to n^4. Since in most practical cases n is quite small, the assumption is well founded.

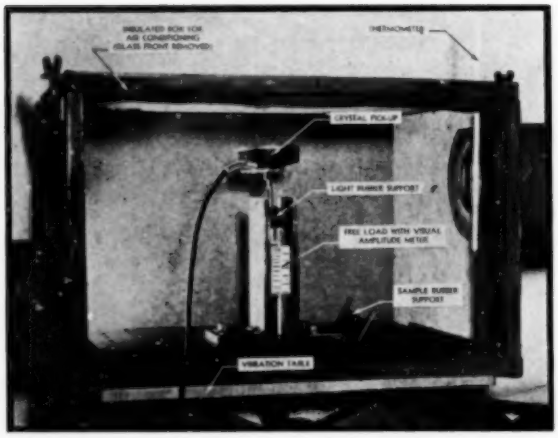


FIG. 4. Small shear type rubber support with 2-lb. load.

or, in other words, the natural frequency

$$\omega_n = p = (c/m)^{\frac{1}{2}}$$

or

$$c = m\omega_n^2. \tag{8}$$

Substituting (7) into Eq. (6) we find:

$$A = \frac{a}{2n}(\omega_n^2 + 4n^2)^{\frac{1}{2}}.$$

Re-substituting:

$$n = \frac{k}{2m},$$

$$A = \frac{a}{k}(\omega_n^2 m^2 + k^2)^{\frac{1}{2}},$$

or

$$k = \frac{\omega_n m a}{(A^2 - a^2)^{\frac{1}{2}}}. \tag{9}$$

Thus, with the observed test data, both c and k can easily be determined.

APPLICATION

Considerable use has been made of the above theory in determining the vibration characteristics of automotive and industrial supports. Figures 2 and 3 show typical examples; one, a light engine support tested under a load of 100 lb., the other (Fig. 3) a heavy motor mounting assembly tested under 400-lb. load. Even heavier loads than this are anticipated, but apparatus capable of handling them has not been completed yet.

Figure 4 shows a very light shear type support, tested under a load of approximately 2 lb. Typical results obtained with this type of support are shown below. A visual amplitude meter is used to locate the point of resonance and is capable of measuring amplitudes to within .01 inch, while a crystal pick-up will give greater accuracies in determining both frequency and amplitude. The output of this pick-up is recorded with an oscillograph.

Figure 5 shows a full view of the light vibration table unit fitted with air conditioning equipment that allows tests to be made at a wide range of temperatures.

TYPICAL RESULTS

The results shown in Table I were gained from a set-up similar to that shown in Fig. 4. Tests were made at room temperature of 25°C.

DISCUSSION

The method described here lends itself well to a speedy determination of dynamic characteristics of rubber supports. The variables of the physical

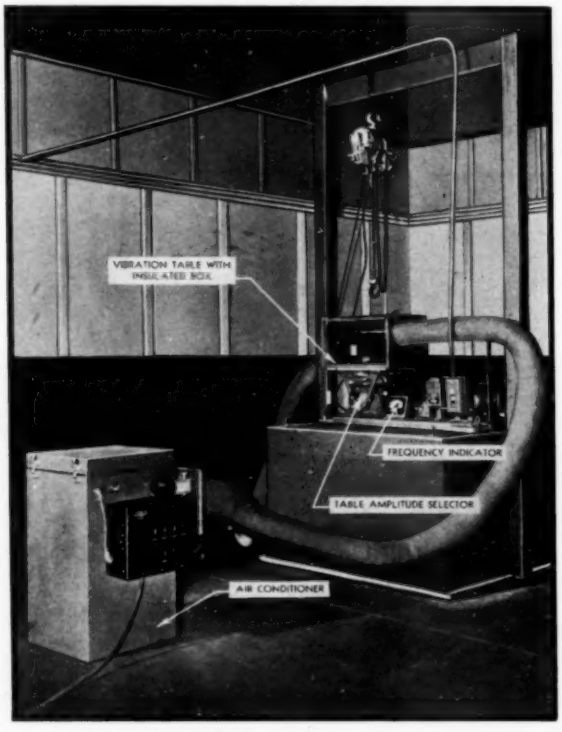


FIG. 5. View of vibration testing equipment.

set-up can be changed quickly and separately to give any desired resonance conditions. They are easily recorded, for displacements, frequencies, and dead weights are the only quantities to be measured. Calculations, also, are simple and straightforward.

A limitation is the fact that it is difficult with the present machines to make tests at constant strain which may, however, be remedied by redesigning the vibration table so that the forcing amplitude can be varied while the machine is running at the resonant frequency of the test piece.

TABLE I.

Type stock	Hevea	Buna-S	Neoprene	Butyl	Buna-N
Shore A durometer	40	41	42	38	40
Load (g)	1027	1027	1027	1027	1027
Nat. freq. (c.p.s.)	23.15	27.85	24.3	26.5	26.2
Dyn. rate (p.s.i.)	70	102	77	92	89
Table ampl. (inches)	.0125	.025	.0125	.025	.025
Dyn. strain	.21	.13	.17	.08	.14
Int. friction (kilopoises)	1.99	7.83	2.63	13.18	6.84

ACKNOWLEDGMENTS

The author wishes to thank the Engineering Research Department of the Firestone Tire and Rubber Company for the equipment used and for permission to publish the photographs. Appreciation is also due Dr. J. H. Dillon, and Messrs. E. F. Riesing, G. L. Hall, and A. B. Lane for their advice and cooperation, Mr. P. P. Crisp for his continued interest, and the Firestone Tire and Rubber Company for permission to publish this article.

The Absorption of Microwaves by Gases

W. D. HERSHBERGER

RCA Laboratories, Princeton, New Jersey

(Received December 15, 1945)

The absorption by ammonia of electromagnetic waves having a length in the one-centimeter range has been known for some years. An investigation has been made to determine whether other gases show similar absorption for microwaves and as a result fourteen additional gases have been found whose absorption is comparable to that of ammonia. Among these gases are dimethyl ether, a variety of amines and alkyl halides, and several others. Measurements on the absorption coefficient and dielectric constant of these gases taken at 1.25 cm at room temperature and a pressure of one atmosphere are given. The frequency at which the absorption coefficient attains its maximum value may be inferred from the curve: absorption coefficient *vs.* pressure. Data on the absorption of several gas mixtures are given. Possible molecular mechanisms adequate to account for the large absorptions observed are discussed together with the conclusions reached.

IT has been known for ten years that ammonia strongly absorbs 1.25-cm microwaves. This paper presents the results of a somewhat detailed study of this effect in ammonia and also gives the results of an investigation which was made to determine whether there are gases other than ammonia which show similar absorption.

Among the absorbing gases found are a variety of alkyl halides, dimethyl ether, and various amines. However, the mechanism of absorption for some of these other gases obviously is different from that involved in ammonia absorption. The order of magnitude of these absorption effects will be made clear by reviewing the pertinent information for ammonia as originally given by Cleeton and Williams.[1] For this gas, at room temperature and at a pressure of one atmosphere, the absorption is such that a plane wave having a wave-length of 1.25 centimeters will be attenuated to $1/e$ of its initial power on traversing a layer of ammonia 1.2 meters thick. Further, the absorption band is relatively wide,

[1] C. E. Cleeton and N. H. Williams, Phys. Rev. **45**, 234 (1934).

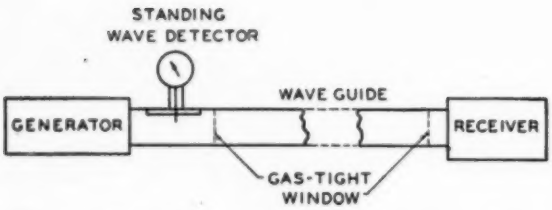


FIG. 1. Wave guide system for measurements on gases.

since under the above conditions absorption falls to one-half of its maximum value at approximately 1.0 centimeter and 1.5 centimeters.

TECHNIQUES

Cleeton and Williams measured absorption in ammonia by passing a parallel beam of microwaves through a cell with a 36-inch by 45-inch cross section such that a transmission path of 16 inches in the gas was traversed. The transmission loss through this cell was observed experimentally on substituting ammonia for air in the cell. Since the time of their investigations, wave-guide measuring techniques[2] have been developed and were used in these experiments. The power absorption coefficient for unconfined plane waves having a length λ in a given gas is given by

$$\alpha_P = \alpha_g[1-(\lambda/2b)^2]^{\frac{1}{2}}, \qquad (1)$$

where λ_g is the experimentally measured absorption coefficient of the gas which is enclosed in a rectangular wave guide used in the H_{01} mode, **b** is the width of the guide.

The derivation of Eq. (1) is given in the appendix as well as that of other quantities that depend on the solution of Maxwell's equations in a wave guide containing a dielectric with small losses.

α_g is determined experimentally by noting an initial power reading P_0 in the receiver and a second power reading P, after introducing the gas under study into a length of guide L situated between the source of microwave power and the receiver.

Then

$$\log P/P_0 = -\alpha_g L,$$

and finally

$$\alpha_P = -([1-(\lambda/2b)^2]^{\frac{1}{2}} \log P/P_0)/L. \qquad (2)$$

[2] G. C. Southworth, Bell Sys. Tech. J. **15**, 284 (1936); W. L. Barrow, Proc. I.R.E. **26**, 1520 (1938); J. C. Slater, *Microwave Transmission* (McGraw-Hill Book Company, Inc., New York, 1942), Chapter III.

Figure 1 shows in schematic fashion the microwave system used in taking measurements. Components used include a generator, a receiver, a standing wave detector, and a section of wave guide. Various guide lengths in the range from three feet to 30 feet were used. Two mica windows sealed into the wave-guide system were used to confine the gas in a given section of the guide. The reflection introduced by a window was compensated by a reflection set up from a tuning element so the over-all effect was a matched guide at the operating frequency. In measurements in which dielectric constant alone was desired, only one window and tuner of this sort were used while the remote end of the guide was closed by a metal diaphragm, thus giving a pronounced standing wave in the system. The position of a minimum in the standing wave pattern was located just outside the window by the use of a standing wave detector. This minimum was located with the test section of guide evacuated, and the shift in its position observed as the gas under study was admitted to the guide. This technique closely resembles that used in optical refractometers in which a specific interference fringe is kept on a cross hair by means of an optical adjustment. The theory of the method as it applies to wave guides is given in the appendix. If ϵ is the specific inductive capacity of the gas and δ is defined by

$$\epsilon = 1+\delta,$$

then

$$\delta = \frac{2\Delta L}{L}\left[1-\left(\frac{\lambda_{vs}}{2b}\right)^2\right], \qquad (3)$$

where ΔL is the shift in the position of the minimum in the standing wave pattern, L is the length of the test section of guide, and λ_{vs} is the free space wave-length.

A method which permits simultaneous determination of specific inductive capacity and absorption coefficient involves the use of a voltage standing wave ratio of the order of 1.1. Under this condition, Fig. 2 shows the wattmeter reading obtained as a function of the pressure of ethyl chloride in a guide 10 feet long. The maxima in this figure occur when the tuning conditions for maximum power transfer have been satisfied. The cyclic variations in wattmeter reading permit calculation of specific inductive

capacity, while absorption is determined from the behavior of the axis of the curve which is shown as a broken line. This method is best suited for work with gases having relatively high specific inductive capacity but only moderate losses. In any event, it is of the utmost importance to have a closely controlled standing wave ratio in measurements by any of the methods outlined.

RESULTS AND DISCUSSION

It was anticipated that the most promising gases for microwave absorption would be those organic gases having a dipole moment and in which there exists the possibility for hindered rotation of a methyl group with respect to some other group about a single carbon-to-carbon bond. The two equilibrium positions in the ammonia molecule for the nitrogen atom are separated by a potential barrier of about 6 kcal. per mole, while the height of the barrier which separates two adjacent minima in the internal rotation energy curve in ethane is estimated to be about 3 kcal. per mole. The energy level splitting responsible for ammonia absorption at 1.25 centimeters decreases exponentially as the square root of the area of the potential hill cut off by an energy level, so it seemed reasonable to expect that the level splitting in ethane-like molecules, while considerably more complicated than in ammonia, would lead to absorption in the microwave frequency region. Accordingly, the first new gas used in taking measurements was ethyl chloride. This gas was found to exhibit about one-fifth the absorption of ammonia.

TABLE I.

Material	$\alpha \times 10^2$	$\delta \times 10^5$
Ammonia	78.	5.5
Methyl fluoride	10.0	9.2
Methyl chloride	8.25	10.5
Methyl bromide	6.6	9.
Ethyl chloride	14.5	12.
Freon 22 $CHClF_2$	10.5	5.4
Freon 21 $CHCl_2F$	10.6	5.8
Freon 12 CCl_2F_2	3.	4.
Dimethyl ether	4.1	4.5
Ethylene oxide	7.	11.
Sulphur dioxide	8.	3.5
Methylamine	9.	4.2
Dimethylamine	7.	3.1
Ethylamine	11.	4.1
Hydrogen sulphide	0.5	2.7

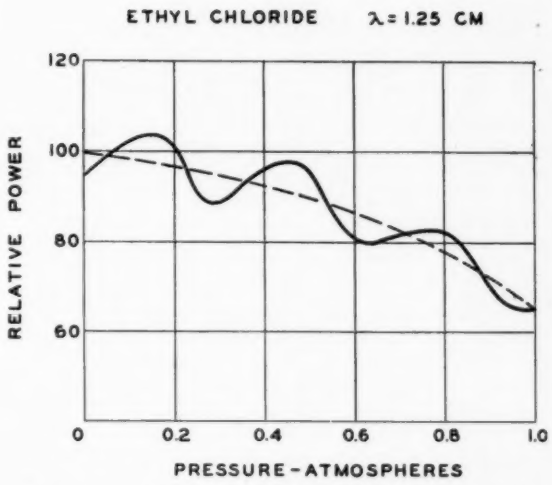


FIG. 2. Receiver power using mismatched wave guide.

Samples of various amines and dimethyl ether were then secured and were also found to absorb microwaves. However, the simple hypothesis involving only hindered rotation was not found to be adequate when measurements showed that the methyl halides and the various gases known commercially as Freon also strongly absorb microwaves. Table I gives the values for absorption and δ for the gases found to be highly absorbing. These values were obtained at room temperature and a pressure of one atmosphere absolute at a wave-length of 1.25 centimeters. The reciprocal of α is the path length in meters which suffices to reduce the power in a plane 1.25-centimeter wave to $1/e$ of its initial value. Thus for ammonia this path length is 1.28 meters while for methyl fluoride it is ten meters. The present measurements indicate that the values reported by Cleeton and Williams for ammonia are perhaps seven percent too high. Measurements on the first ten materials were taken with the 30-foot guide and are accordingly more accurate than the measurements on the remaining five substances where the 5-foot section only was used.

Polyatomic gases whose absorption is either zero or too small for measurement with the present technique include: hydrogen, oxygen, nitrogen, carbon dioxide, carbon monoxide, ethylene, acetylene, nitrous oxide, methane, ethane, propane, propylene, *n*-butane, isobutane, butene-1, butene-2, isobutylene, butadiene, and trimethylamine.

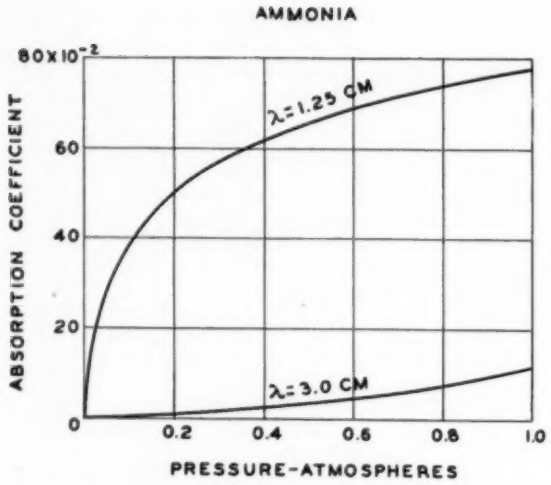


FIG. 3. Absorption of ammonia at 1.25 and 3.0 centimeters.

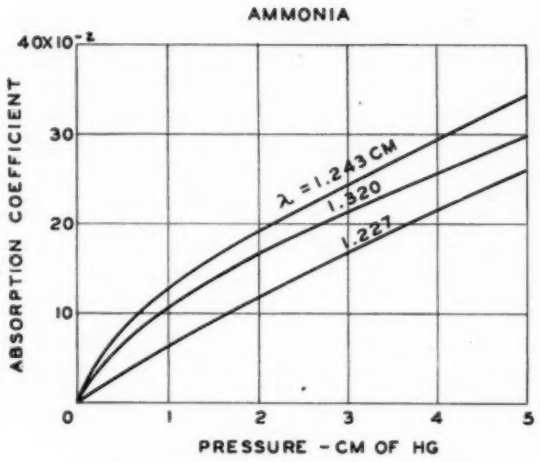


FIG. 4. Absorption of ammonia at 1.227, 1.243, and 1.320 centimeters.

The smallest absorption which it is possible to detect using a 30-foot section of wave guide is in the neighborhood of 0.5×10^{-2} per meter, which means that measured power at the receiver falls to 94 percent of its initial value on introducing a gas characterized by this absorption coefficient. This is the same change that one would observe in a mismatched wave guide system with a voltage standing wave ratio of 1.03 and which arises because the specific inductive capacity of the gas differs from unity. Thus the problem reduces to one of distinguishing between monotonic received power changes which arise from gas losses, and the cyclic power changes which arise purely from wave guide mismatch.

For low loss materials, δ is less than 10^{-3} and may be as low as 10^{-4}. It will be noted that in Eq. (3) the factor $2[1-(\lambda_{vs}/2b)^2]$ is of the order of unity, so δ approximately equals $\Delta L/L$. Thus if L is three meters, and δ of the order of 10^{-3}, we must measure ΔL to 3 millimeters, which is feasible. However, if we attempt to increase precision by increasing L to, say, six meters, the standing wave pattern becomes poorly defined owing to copper losses in the guide giving rise to considerable uncertainty in setting the standing wave detector, hence giving us no marked increase in precision.

The location of the maximum in the curve relating absorption and frequency is of great theoretical interest in studies involving molecular structure. The maximum in this curve was determined for ammonia by Cleeton and Williams by the use of a series of microwave generators and they made a point-by-point experimental determination. A useful method entailing less labor is to employ experimentally determined curves relating absorption and pressure taken at several discrete frequencies and then use the theory of Lorentz broadening of spectral lines to determine the parameters which enable this curve to be constructed. The character of these pressure-absorption coefficient curves for ammonia is shown in Figs. 3 and 4. When the operating frequency is nearly the same as the proper molecular frequency, the curve displays a knee at a relatively low pressure. The effect may be described qualitatively as follows: At first absorption increases very nearly linearly with pressure until a pressure is reached such that collision frequency in the gas becomes comparable with the operating frequency. At still higher pressures, the increase in absorption due to increase in density is very largely offset by the fact that intermolecular collisions limit the length of the absorbed wave trains. The spectral distribution in these interrupted wave trains is such that very little absorption is gained at the proper molecular frequency but absorption at neighboring frequencies only is enhanced. At a frequency well removed from the proper frequency, absorption may increase with the square of the pressure. The curve for 3.0 centimeters in Fig. 3 shows an increase in absorption of this character. The curves in Fig. 4 show how sen-

sitive absorption is to frequency at reduced pressures.

The effect on absorption of impurities was studied by mixing a non-absorbing gas with an absorbing gas, and some results are shown in Fig. 5. Mixtures of ammonia with hydrogen, with oxygen, and with methane were used and no difference in the net effect could be measured which might arise because of the quite wide range in molecular weights of the non-absorbing gases—from two to 32—in comparison to the molecular weight of the ammonia which is 17.03. The solid lines represent the experimental results obtained. The dashed lines which were drawn in later represent the absorption of a constant quantity of ammonia in a given length of wave guide but with an increasing partial pressure of the non-absorbing gas. The curves indicate that the non-absorbing gas has a relatively small yet measurable effect on the absorption of the ammonia; that is, absorption is to a first approximation dependent on the partial pressure of the ammonia. The presence of the foreign gas presumably widens the absorption band at the expense of absorption at its maximum. In these experiments it is essential to use premixed gases; that is, quite misleading results will be obtained if ammonia is admitted to an evacuated guide to some chosen pressure, and the foreign gas is then admitted at one end of a guide section compressing the ammonia before it into one end of the section. Under these conditions the manner in which net absorption changes with time might well be made the basis of measurements in diffusion rates in slender tubes.

Absorption by the methyl halides strongly suggests that we are dealing with transitions between the lowest rotational states of these molecules. However, the calculated absorption due to rotation for these gases is much less than that observed. The theoretical absorption for methyl chloride is approximately the same as that calculated for carbonyl sulphide which is triatomic and linear. The use of this material in a critical check on the theory was suggested by J. Turkevich. However, measurements show that the absorption of carbonyl sulphide is certainly less than one-tenth that of methyl chloride. The small calculated absorption due to rotation arises primarily from the sparsity of the molecular populations in the lower rotational states at room temperature. Thus but one molecule in 1000 of carbonyl sulphide is in the appropriate rotational state for 1.25-cm absorption at room temperature. Hence, we are forced to look to "inversion doubling" of energy levels in the methyl halides akin to that found in ammonia to account for the experimental results. Further studies designed to evaluate the contributions from all causes to the measured absorption are now under way as well as analyses of pressure *versus* absorption curves for the materials listed in Table I.

APPENDIX: THE EFFECT OF THE DIELECTRIC ON PROPAGATION IN WAVE GUIDES

The H_{01} or TE_{10} solution of Maxwell's equations which satisfies the boundary conditions imposed by the conducting walls of a rectangular wave guide of width **b** has the form

$$E_x = A \sin(\pi y/b) \exp(j\omega t - \nu z),$$
$$H_y = B \sin(\pi y/b) \exp(j\omega t - \nu z),$$
$$H_z = H_0 \cos(\pi y/b) \exp(j\omega t - \nu z).$$

It is, of course, trivial to express A and B in terms of H_0 by the use of Maxwell's equations. However, here the interest centers in the propagation function ν which is complex when we have losses in the dielectric. Each of the components of **H** or **E** satisfies the wave equation which now includes a dissipation term. We require ν in terms of wave guide dimensions, operating fre-

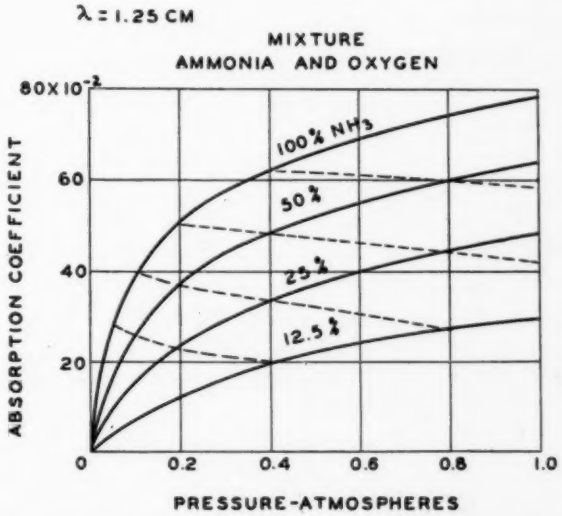


FIG. 5. Absorption of mixtures of ammonia and oxygen.

quency $f=\omega/2\pi$, and the properties of the dielectric. Thus H_z must satisfy the wave equation

$$\nabla^2 H_z-\sigma\mu\dot{H}_z-\mu\epsilon\ddot{H}_z=0.$$

This is true if

$$\nu^2=(\pi/b)^2-\omega^2\mu\epsilon+j\omega\mu\sigma.$$

If we let $\nu=a+j\beta$,

$$\beta^2=\omega^2\mu\epsilon-(\pi/b)^2+a^2,$$

while

$$2a=\omega\mu\sigma/\beta.$$

β determines wave-length in the guide and the speed with which power is transmitted, which **a** is the amplitude damping factor. Hence **2a** is the power absorption coefficient. ϵ is the permittivity of the dielectric inside the guide, μ is the permeability of this dielectric while σ is its conductivity.

In this paper the m.k.s. system of units is employed. Thus for free space, we use the subscript zero and have $\epsilon_0=8.85$ micromicrofarads per meter, $\mu_0=1.257$ microhenries per meter, while

$$c=[\mu_0\epsilon_0]^{-\frac{1}{2}}.$$

For media such as those considered here

$$\epsilon=\mathfrak{e}\epsilon_0, \quad \mu=\mathfrak{u}\mu_0,$$

where $\mathfrak{e}$ is the specific inductive capacity and $\mathfrak{u}$ is the specific permeability. For attenuation of the order of magnitude encountered in gases, **a** has a negligible effect on wave-length so

$$\beta=(\omega/v)[1-(\lambda/2b)^2]^{\frac{1}{2}},$$

where $v=[\mu\epsilon]^{-\frac{1}{2}}$ which is the speed of propagation in the gas for unconfined waves. Also

$$\lambda=v/f,$$

while $\lambda_0=2b$ is the cut-off wave-length for the guide; thus there is no wave propagation when $\lambda>\lambda_0$. Then

$$\alpha_g=2a=\sigma(\mu/\epsilon)^{\frac{1}{2}}[1-(\lambda/2b)^2]^{-\frac{1}{2}}.$$

But $\sigma(\mu/\epsilon)^{\frac{1}{2}}$ is the power absorption coefficient α_P for plane waves. Or

$$\alpha_P=\alpha_g[1-(\lambda/2b)^2]^{\frac{1}{2}}.$$

This is Eq. (1).

To derive Eq. (3) dealing with the shift in the position of the minimum in the standing wave pattern resulting from the admission of gas into a wave guide, we take $\mathfrak{u}=1$ and again neglect the effect of attenuation on wave-length in the guide.

Then

$$\beta^2=\mathfrak{e}(\omega/c)^2-(\pi/b)^2.$$

Wave-length in the guide is given by

$$\lambda_g=2\pi/\beta.$$

Let λ_{vs} represent the free space wave-length in vacuum. Then

$$\lambda_g=\lambda_{vs}[\mathfrak{e}-(\lambda_{vs}/2b)^2]^{-\frac{1}{2}}.$$

We let

$$\mathfrak{e}=1+\delta,$$

and since the wave-length in the guide in vacuum is given by

$$\lambda_{vg}=\lambda_{vs}[1-(\lambda_{vs}/2b)^2]^{-\frac{1}{2}},$$

we have finally

$$\lambda_g/\lambda_{vg}=\left\{1-\frac{\delta}{2}[1-(\lambda_{vs}/2b)^2]^{-1}\right\},$$

treating δ as a quantity which is small compared to unity. ΔL is the shift measured in this investigation and the length of guide containing the gas is L. Let n be the number of the wave-lengths in the guide when evacuated. Then

$$n\lambda_{vg}=L.$$

We now admit gas to the guide but vary the position of the probe in such a manner as to follow a specific minimum in the standing wave pattern. Then

$$n\lambda_g=L-\Delta L,$$

and

$$\lambda_g/\lambda_{vg}=1-\Delta L/L.$$

Hence

$$\delta=\frac{2\Delta L}{L}[1-(\lambda_{vs}/2b)^2]$$

which is Eq. (3).

When δ is so large that the approximation used in deriving (3) is not justified, we can always go back to the exact expression, namely

$$(\lambda_{vs}/\lambda_g)^2+(\lambda_{vs}/2b)^2=\mathfrak{e}.$$

Pressure-Volume Relation for Cylindrical Tubes with Elastomeric Walls: The Human Aorta

ALLEN L. KING
Dartmouth College, Hanover, New Hampshire
(Received December 19, 1945)

The general theory of the elasticity of elastomers as developed by James and Guth (reference 9) is extended and applied to cylindrical elastomeric tubes. Data obtained by Hallock and Benson (reference 4) on human aortas of several age groups are described rather well by the resulting equations. On the basis of this analysis an interpretation of the aging process for aortas is given.

DURING systole, blood from the left ventricle of the heart is forced into the aorta. From here it is distributed throughout the body by means of systemic arteries. With their elastic walls, the aorta and large systemic arteries apparently regulate the flow of blood to the capillary bed by smoothing the pulsating outflow from the heart—much as an electrical filter smooths out the pulsating d.c. from a rectifier circuit. With advancing age the arterial walls become less elastic and thicker as collagenous fibers increase in number and elastic tissue seemingly disappears. Lesions often occur, especially in old age, and senile degeneracy sets in. In spite of these several changes, the aorta and immediate systemic arteries continue to serve as a regulating system.[1,2]

Evidently, elastic arteries have an important function in the cardiovascular system. At the same time, they are subject to an irreversible aging process which decreases their efficiency in smoothing out the pulsating blood flow. Many investigators have measured elastic characteristics of arterial walls in the hope that some insight would be gained regarding the nature of this aging process.

In 1880 Roy[3] published an extensive report on the elastic properties of the arterial wall. He showed that animal tissues have thermoelastic properties comparable to those of caoutchouc and established the fact that static pressure-volume curves for human aortas are normally S-shaped, especially in youth. In 1937 Hallock and Benson[4] published measurements on segments of human aortas obtained post-mortem for several age groups and showed how the pressure-volume curve is changed in shape with advancing age. Observations on strips of aortas yield similar results.[5] In all cases, normal healthy aortas have been found to possess non-linear characteristics and a decidedly rubber-like behavior. For convenience, therefore, in the following discussion an arterial wall will be considered replaced by a uniformly thick cylindrical *elastomeric* shell.

The aortic wall has been found to consist of three fairly well-differentiated layers.[6] Its great extensibility is ascribed to the presence of elastic tissue in all three layers, but especially in the *tunica medialis*. The principal constituent of elastic tissue is an albuminoid called elastin. Interspersed with elastin are found other relatively inelastic substances. As a matter of fact, even these latter substances probably should be considered elastomeric. Because of their presence, the actual aortic wall has a greater thickness than the ideal one to be discussed below.

As mentioned above, an important factor generally recognized as influencing the rate of blood flow in the aorta is the distension of the wall. Elsewhere[7] a general relation for the rate of flow of viscous fluids in cylindrical tubes with distensible walls has been given. In order to

[1] C. J. Wiggers, *Physiology in Health and Disease* (Lea and Febiger, Philadelphia, 1944), p. 555; H. D. Green, *Glasser's Medical Physics* (Year Book Publishers, Chicago, 1944), p. 208.
[2] E. V. Cowdry (ed.), *Arteriosclerosis* (The Macmillan Company, New York, 1933).
[3] C. S. Roy, Foster's J. Physiol. **3**, 125 (1880).
[4] P. Hallock and I. C. Benson, J. Clinical Invest. **16**, 595 (1937).
[5] J. Krafka, Arch. Pathology **29**, 303 (1940).
[6] Reference 2, pp. 53–61.
[7] A. L. King, A. P. S. Bull. **20**, 8 (1945).

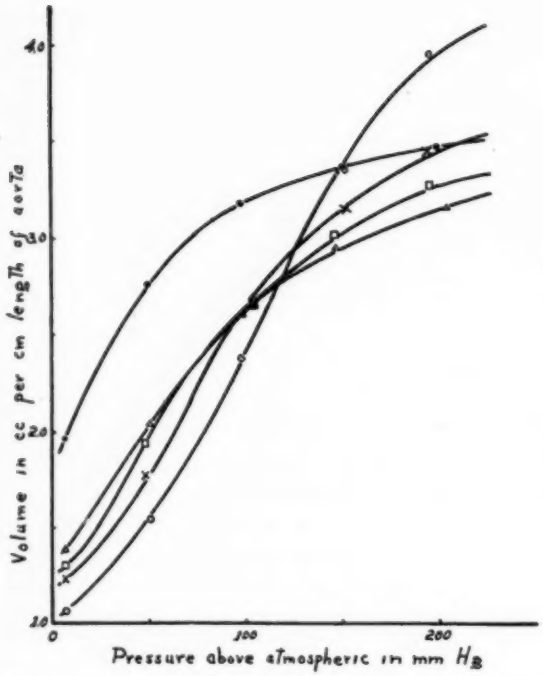


FIG. 1. Curves show Hallock and Benson's observed volumes per unit length of aorta as a function of applied fluid pressure. Theoretical values are plotted at five points for each age group: ○ 20–24 yr., × 29–31 yr., □ 36–42 yr., △ 47–52 yr., ● 71–78 yr.

obtain an explicit function for this rate, the variation of pressure with radius along the tube needs to be known. Although this variation should be a dynamic one, the static pressure-radius relation deduced below may be used to a first approximation.

During the preparation of this article, Rashevsky[8] published in his journal an interesting treatment of the blood-flow problem. In his analysis he assumes the arterial wall to follow Hooke's law, although several well-known experimental results, such as those mentioned above, suggest that an entirely different approach should be made.

In this article a pressure-volume relation for cylindrical tubes with elastomeric walls is derived and applied to Hallock and Benson's data on human aortas.

THE PRESSURE—VOLUME RELATION

From the general theory of elastomers, the force exerted at right angles to a surface yz by a mesh of molecular chains is given by[9,10]

$$F_x = B_x y_0 z_0 \mathfrak{L}^{-1}(k_x x), \tag{1}$$

where $y_0 z_0$ represents the undistorted area and $\mathfrak{L}^{-1}(k_x x)$ is the inverse Langevin function of argument $k_x x$. k_x is a parameter which depends inversely on the maximum lengths of the molecular chains. If t represents the inverse Langevin function, then

$$\mathfrak{L}(t) = \coth t - 1/t \tag{2}$$

is the Langevin function, where $t = F_x/B_x y_0 z_0$ and $\mathfrak{L}(t) = k_x x$. The coefficient B_x has been assigned various values. For this discussion, however, the exact value does not enter. Nevertheless, it is of some interest to note that it increases directly with absolute temperature.

Consider a cylindrical tube of length z_0 with an elastomeric wall of thickness e_0 and internal

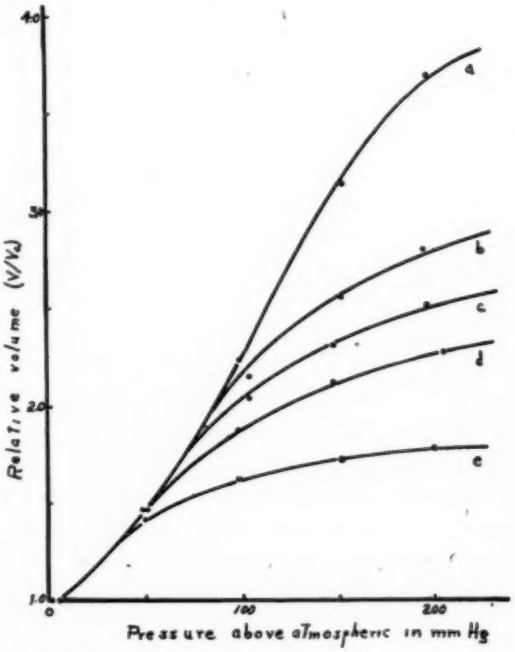


FIG. 2. Relative volumes of aortas in five age groups as a function of applied pressure. The curves are obtained from experimental data and the points from theory. (a) 20–24 yr., (b) 29–31 yr., (c) 36–42 yr., (d) 47–52 yr., (e) 71–78 yr.

[8] N. Rashevsky, Math. Biophys. **7**, 25, 35 (1945); see also O. Frank, Zeits. f. Biol. **71**, 255 (1920).

[9] F. T. Wall, J. Chem. Phys. **10**, 485 (1942); **11**, 527 (1943). H. M. James and E. Guth, J. Chem. Phys. **11**, 455 (1943). See also A. V. Tobolsky and R. D. Andrews, J. Chem. Phys. **13**, 3 (1945). E. Guth, "The Problem of the Elasticity of Rubber and Rubber-like Materials" in *Surface Chemistry*, F. R. Moulton (ed.) (A. A. A. S., Smithsonian Inst., Washington, D. C., 1943), pp. 103–127.

[10] A. L. King, Am. J. Phys. **14**, 30 (1946).

radius r_0, oriented so that its axis lies along the z coordinate of a set of cylindrical polar axes. Initially, the pressure within and outside the tube is assumed to equal p_0. For the moment, the weight of the tube is considered sufficiently small so that any distortion of the tube due to it may be neglected. As a further simplification, the thickness e_0 is assumed to be small compared with r_0. Even for e_0 as much as 20 percent of r_0 this approximation appears justified, as suggested later in the application of this theory to human aortas.

Experimental[11] and theoretical[12] evidence indicates that the volume of a rubber-like material should remain constant as it is stretched up to 300 percent elongation. Thus, if the radius of the elastomeric wall changes to r due to an increased pressure within the tube, the thickness and length change to e and z, respectively, so that

$$2\pi rez = 2\pi r_0 e_0 z_0. \tag{3}$$

For the present problem suppose e_0 and z_0 are multiplied by equal factors; then it follows that for Eq. (3) to be satisfied,

$$e = e_0(r_0/r)^{\frac{1}{2}} \quad \text{and} \quad z = z_0(r_0/r)^{\frac{1}{2}}. \tag{4}$$

In order to find the elastic forces within the wall, the randomly twisted and intermeshed molecular chains are replaced by a three-dimensional network of chains extending along the tube wall, around the tube wall, and radially between the outer and inner surfaces of the wall. These chains usually exert a tension on their colinear neighbors or on any free surface in which they end, since the most probable state for them is one of more complete coiling and random twisting. On the other hand, due to their sidewise thermal motions, the chains push outward on a free surface until the contractile, thermal, and any external forces thereon are in equilibrium. The thermal force may be considered to equal a uniform internal pressure P multiplied by the area of the free surface.

Since the molecular chains around the tube wall do not end on free surfaces, the internal peripheral force F_c comes out to be

$$F_c = e_0 z_0 B_c \mathfrak{L}^{-1}(2\pi r k_c). \tag{5}$$

The longitudinal force is

$$F_l = 2\pi r_0 e_0 B_l \mathfrak{L}^{-1}(k_l z) - 2\pi reP; \tag{6}$$

and the radial force is

$$F_r = 2\pi r_0 z_0 B_r \mathfrak{L}^{-1}(k_r e) - 2\pi rzP. \tag{7}$$

Pressure P may be found from the fact that F_r must equal $2\pi rzp_0$, since the outside surface of the wall is in equilibrium. Furthermore, the ends of the tube are supposed to be acted on by an external force Z which can become adjusted so that at all times $F_l + Z = 0$.

To obtain a relation between pressure p and radius r, a small increase δr in the radius is assumed to be accompanied by incremental changes δe and δz in thickness and length, respectively. Then, by the principle of virtual work:

$$2\pi rz(p - p_0)\cdot\delta r - Z\cdot\delta z = 2\pi F_c\cdot\delta r + F_r\cdot\delta e + F_l\cdot\delta z,$$

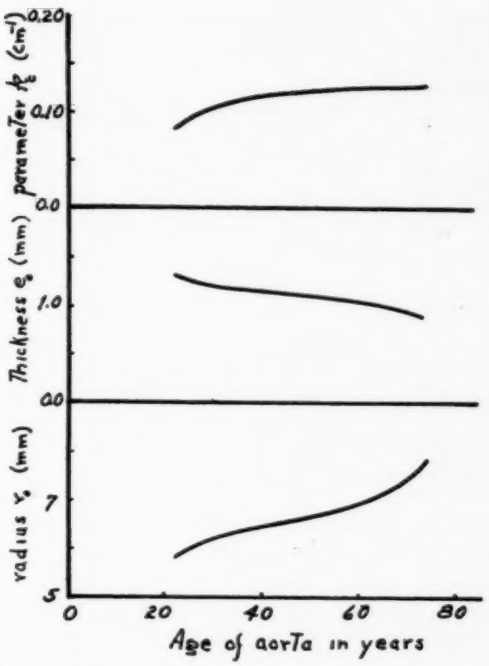


FIG. 3. Variation of radius, wall thickness, and a characteristic parameter with age of undistended aortas. The wall thickness here is an effective value. The parameter depends on the reciprocal length of molecular chains within the aortic wall.

[11] W. L. Holt and A. T. McPherson, J. Research Nat. Bur. Stand. **17**, 657 (1936).
[12] D. R. Elliot and S. A. Lippmann, J. App. Phys. **16**, 50 (1945).

TABLE I. Constants for human aortas of various ages.

Age groups years	V_0 cc	r_0 mm	β	A mm Hg	e_0 mm	ew_0 mm²	k_c cm⁻¹	d_0 cm
20–24	1.07	5.8	0.302	88	1.33	7.7	0.083	1.41
29–31	1.23	6.2	.415	74	1.20	7.45	.106	1.48
36–42	1.30	6.4	.462	70	1.17	7.5	.114	1.52
47–52	1.39	6.6	.510	65	1.12	7.4	.122	1.54
71–78	1.96	7.9	.640	44.5	0.92	7.3	.129	1.76

so that

$$p-p_0=B_c\frac{e_0z_0}{rz}\mathfrak{L}^{-1}(2\pi rk_c)+p_0(\delta e/\delta r). \quad (8)$$

But from Eq. (4), $\delta e=-\frac{1}{2}(e_0/r_0)(r_0/r)^{\frac{1}{2}}\cdot\delta r$ and letting $\beta=2\pi r_0k_c$, Eq. (8) reduces to

$$p-p_0=(e_0/r_0)[B_c(r_0/r)^{\frac{1}{2}}\mathfrak{L}^{-1}(\beta r/r_0)-\tfrac{1}{2}p_0(r_0/r)^{\frac{1}{2}}]. \quad (9)$$

Since $p=p_0$ when $r=r_0$, evidently $p_0=2B_c\mathfrak{L}^{-1}(\beta)$. Thus B_c may be eliminated from Eq. (9) to get the desired pressure-radius relation

$$p-p_0=A\left[(r_0/r)^{\frac{1}{2}}\frac{\mathfrak{L}^{-1}(\beta r/r_0)}{\mathfrak{L}^{-1}(\beta)}-(r_0/r)^{\frac{1}{2}}\right], \quad (10)$$

where $A=e_0p_0/2r_0$.

The initial volume within the tube is $V_0=\pi r_0^2z_0$ and the final volume is $V=\pi r^2z$, so that $V/V_0=(r/r_0)^{\frac{1}{2}}$; and the desired pressure-volume relation for cylindrical tubes with elastomeric walls is found to be

$$p-p_0=A\left[(V_0/V)^{\frac{1}{2}}\frac{\mathfrak{L}^{-1}\{\beta(V/V_0)^{\frac{1}{2}}\}}{\mathfrak{L}^{-1}(\beta)}-(V_0/V)\right]. \quad (11)$$

Because k_c, which depends on the reciprocal of the maximum molecular chain lengths, cannot be determined readily, it is convenient to consider $\beta(=2\pi r_0k_c)$ an adjustable parameter of Eq. (11). Pressure p_0 is that of the surrounding medium; for the above analysis it is assumed constant. If the surrounding medium exerts a variable pressure, however, Eq. (11) can be modified accordingly.

Frequently the weight of the tube, and possibly other small extraneous factors, may cause a partial collapse of an elastomeric tube unless the pressure within is somewhat in excess of that outside. This means the actual external pressure should be considered replaced by a somewhat greater effective pressure.

THE HUMAN AORTA

Perhaps the most widely quoted recent experimental data on human aortas are those of Hallock and Benson. Isolated segments, eight to ten centimeters long, of the descending aorta from just beyond the arch were obtained in the age groups 20–24, 29–31, 36–42, 47–52, and 71–78 years. As far as could be determined all specimens but those of the last group were healthy and normal. Those beyond seventy years showed some degree of coronary disease. The experimental curves are shown in Fig. 1.

To prevent collapse of the aortas an initial pressure of 10-cm saline solution ($\sim$7.5 mm Hg) was used in every case. The volumes per unit length for this initial pressure are given in Table I. Pressures were measured relative to atmosphere by means of a manometer. Since Hallock and Benson fail to state the prevailing atmospheric pressure for any of their observations, in the following discussion the standard value of 760 mm Hg is assumed to apply in all cases. But 7.5-mm Hg pressure are required to balance weight sag and other small factors, so that the effective external pressure p_0 is taken to be 768 mm Hg.

A better appreciation of the relative variation in pressure-volume characteristics with age may be obtained by plotting V/V_0 as a function of pressure difference as shown in Fig. 2. These curves are well represented by Eq. (11) of the above theoretical discussion. For convenience both β and A were adjusted so as to fit these curves. The resulting values are given in Table I. *The function is very sensitive to slight variations in β.*

Since V_0 represents the initial volume per unit length of aorta, the initial radius r_0 is given by $(V_0/\pi)^{\frac{1}{2}}$. Values of r_0 are listed in Table I and are plotted in Fig. 3. With these values of A and r_0 and $p_0=768$ mm Hg, the effective initial wall thicknesses e_0 are computed from $e_0=2r_0A/p_0$ and also are plotted in Fig. 3. Notice that e_0 decreases with advancing age; probably this decrease is associated with the apparent disappearance of elastic tissue. On the other hand,

the product $e_0 r_0$, as seen in Table I, is constant, at least for the data under examination. The mean value is 7.5 mm² with an average deviation of only ±0.1. As dilation of an aorta proceeds with advancing age, the *effective* cross-sectional area of the tube wall remains nearly constant and equal to 47 mm² between the ages of 20 and 80 years. Perhaps this observation may be interpreted as follows: the materials within the vessel wall that render it elastic really do not disappear, but rather become more thinly distributed in the enlarged wall of the older aortas. More extensive experimental data, however, may show a real divergence from the above mean, especially at either end of the time scale, in one case due to new tissue growth and in the other to senile alterations.

External diameters d_0 of the undistended aortas approximately equal $2(r_0+e_0)$. The values shown in Table I are in rather good agreement with measured values. For instance, Roy[13] observed circumferences of 45½ mm for a 22-year old and 52 and 55 mm for 71- and 76-year old aortas. These values correspond to diameters of 1.4 cm for the young aorta and 1.7 cm for the old aortas. Observed wall thicknesses, on the other hand, vary greatly, as illustrated by Krafka's data for *strips* of human aortas having no marked sclerotic lesions (Table II).[5] These values, for the most part, are outside the limits of 1–2 mm mentioned by Krafka in an earlier paper.[14] It should be recognized, however, that on cutting strips from an aorta, peripheral chains are shortened and made to end on free surfaces. The equilibrium is thereby altered and, perhaps, the wall then becomes thicker. In any case, as suggested previously, e_0 should be expected to have values somewhat less than the actual wall thicknesses.

TABLE II. Observed wall thicknesses of human aortas.

Age (years)	7	20	31	37	37	41	47	48
Thickness (mm)	1.8	2.6	2.6	2.9	1.76	3.18	2.34	2.46

From $\beta = 2\pi r_0 k_e$ the parameter k_e is computed and entered in Table I. Its variation with age is shown in Fig. 3. Evidently k_e is an important age-dependent physical characteristic of aortas. Since it is defined as the molecular chain length reciprocal, clearly its increase with age corresponds to an effective decrease in distance between chain ends around the aorta. Perhaps a not unreasonable interpretation may be this: with advancing age, bonds between chains become established, thus making the effective chain lengths less and, consequently, producing a more rigid structure. In other words, arterial walls may be thought to lose their elasticity as a result of increased interchain bonding. Notice that chain lengths are reduced, on the average, by a factor of nearly two-thirds during the age interval from 20 to 80 years. The parameter β itself is a factor which represents the amount of coiling and twisting experienced by the peripheral molecular chains. Thus the distance between ends of a single chain around the aorta is only 0.302 of the maximum chain length for 20–24 year old specimens, whereas this ratio becomes 0.640 for 71–78 year old specimens.

The above analysis does not exclude the possibility of chains aggregated in strands. Perhaps it is such strands which produce the wavy dark streaks in sections of aortic wall that have been properly stained to show the presence of elastin.

[13] C. S. Roy (reference 3), pp. 148 and 152–153.
[14] J. Krafka, Am. J. Physiol. **125**, 5 (1939).

Approximations to Planckian Distributions

PARRY MOON
Massachusetts Institute of Technology, Cambridge, Massachusetts
AND
DOMINA EBERLE SPENCER
Tufts College, Medford, Massachusetts
(Received January 22, 1946)

The use of Planck's equation in theoretical photometry and colorimetry has been hampered by the difficulty of integration. The following investigation attempts to eliminate this obstacle by replacing Planck's equation in the visible region by simpler expressions that are easily integrated. Two such approximations are employed—a three-term polynomial and the Wien function. Both allow analytic integrations to be performed with considerable ease and satisfactory accuracy. The method applies even in the weighted integrations of colorimetry: the standard weighting functions representing characteristics of the human eye are introduced in their analytic form and make possible the use of analytic methods in place of the unwieldy numerical approximations employed previously.

1. INTRODUCTION

THE importance of Planck's equation in photometry and colorimetry can hardly be overestimated. Unfortunately, however, the equation does not lend itself to easy manipulation. Ratios of Planck functions obstinately keep their form instead of simplifying as Wien functions do, and integrations are performed only with difficulty.

It seems doubtful if a simple approximation to the Planck equation is possible when the entire wave-length scale is needed. But in photometry and colorimetry, only the visible region is considered, and in this region the radiation curves have very simple shapes which can be approximated by elementary functions.

The following pages present the results of an investigation of approximations to Planckian distributions. Both polynomials and Wien functions are employed. In either case, analytic integrations are easily performed and various kinds of mathematical manipulation are expedited.

TABLE I. Polynomial approximation to Planckian distributions. $J(\lambda)=K_0+K_1\lambda+K_2\lambda^2$ (watt m^{-2} $micron^{-1}$.)

T (°K)	K_0	K_1	K_2
2000	2.15635×10^{5}	-9.7454×10^{5}	11.1290×10^{5}
2100	3.03280	−14.0909	16.5959
2200	3.95017	−19.0373	23.3151
2300	4.73990	−24.0395	30.9838
2400	5.15858	−28.2523	39.0340
2500	4.88558×10^{5}	-30.5080×10^{5}	46.6020×10^{5}
2600	3.53338	−29.3414	52.5395
2700	+ 0.65957	−23.0210	55.4300
2800	− 4.21592	− 9.6010	53.6250
2842	− 6.97760	− 1.4168	51.0905
2900	−11.57840	+12.9650	45.3400
3000	-2.19248×10^{6}	$+4.6864\times10^{6}$	2.8599×10^{6}
3100	− 3.56774	9.4090	+ 0.1550
3200	− 5.32508	15.6716	− 3.7805
3300	− 7.49365	23.6450	− 9.1200
3400	−10.09439	33.4790	−16.0250
3500	-1.31403×10^{7}	4.5300×10^{7}	-2.4642×10^{7}
4000	− 3.43851	13.5166	− 9.5855
4500	− 6.04611	26.7200	−21.1150
5000	-7.84826×10^{7}	40.8458×10^{7}	-34.9482×10^{7}
5500	− 7.13490	50.9390	−47.4900
6000	− 2.10080	51.4390	−54.4858
6500	+ 8.89830	37.0600	−51.7500
7000	$+2.71395\times10^{8}$	3.5735×10^{8}	-3.5870×10^{8}
20000	$+3.09456\times10^{10}$	-8.4853×10^{10}	$+6.1053\times10^{10}$

2. POLYNOMIAL APPROXIMATION

The simplicity of the plot of Planck's equation in the visible region suggests the possibility of approximating it by a few terms of a power series. At temperatures above 2500°K, a fairly satisfactory representation is obtained with two terms:

$$J(\lambda)=K_0+K_2\lambda^2, \tag{1}$$

where K_0 and K_2 are constants and λ is the wavelength. At 2000°K, however, three terms are found to be necessary; and in the interests of standardization, three terms were adopted for approximating Planckian radiation at all temperatures.

The equation being approximated is

$$J(\lambda)=\frac{C_1}{\lambda^5}\frac{1}{\exp(C_2/\lambda T)-1}, \tag{2}$$

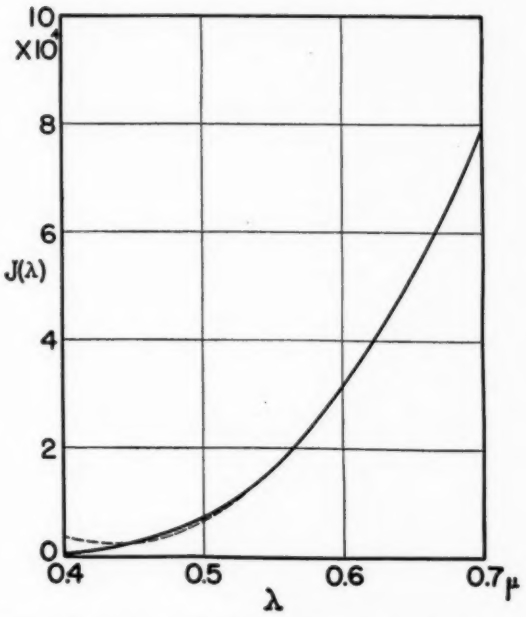


FIG. 1. Planckian radiation at 2000°K, showing the kind of fit that is obtained by the polynomial approximation at low temperatures. Ordinates are in watt m^{-2} $micron^{-1}$ at the radiator opening. ——— Planckian radiation; - - - - three-term analytic approximation.

where[1] $C_1=3.6970\times10^8$, $C_2=14320$, λ=wave-length (micron), and $J(\lambda)$=spectral pharosage (watt m^{-2} $micron^{-1}$ at blackbody opening). The approximate equation is

$$J(\lambda)=K_0+K_1\lambda+K_2\lambda^2. \qquad (3)$$

If $J(\lambda)$ from Eq. (3) is made to coincide with $J(\lambda)$ from Eq. (2) at wave-lengths 0.45, 0.55, and 0.65μ, the constants K_0, K_1, and K_2 are determined by the equations,

$$\left.\begin{aligned} K_0&=17.875J(0.45)-29.250J(0.55)\\ &\qquad +12.375J(0.65),\\ K_1&=-60J(0.45)+110J(0.55)\\ &\qquad -50J(0.65),\\ K_2&=50J(0.45)-100J(0.55)+50J(0.65). \end{aligned}\right\} \qquad (4)$$

The constants of Eq. (3) are evaluated by means of Eq. (4) for a number of temperatures, and the results are listed in Table I. Values for other temperatures are easily computed by means of Eqs. (2) and (4). The discrepancies between Planckian radiation and the three-term approximation may be visualized by considering Figs. 1 to 4. These graphs cover the range ordinarily employed in colorimetry and show that a reasonably good fit is obtained at all temperatures through most of the visible region. The discrepancies at the ends of the visible region do not prove troublesome in most applications. It is interesting to note that curves of such diverse shapes as Figs. 2 and 4 can be fitted by the same form of expression.

If a more accurate representation is desired, a seven-term polynomial may be used:

$$J(\lambda)=K_0+K_1\lambda+K_2\lambda^2+\cdots+K_6\lambda^6. \qquad (5)$$

The details in the evaluation of the constants are discussed elsewhere.[2] For example, take Planckian radiation at 2842°K and pass the approximate curve through the correct values at 0.40, 0.45, 0.50, $\cdots$0.70μ. The constants are found to be,

$$\begin{aligned} K_0&=-1{,}185.36370,\\ K_1&=+13{,}764.24041,\\ K_2&=-63{,}762.05721,\\ K_3&=+148{,}682.59032,\\ K_4&=-182{,}496.89139,\\ K_5&=+114{,}098.58438,\\ K_6&=-28{,}888.87725. \end{aligned}$$

A comparison of the true values of $J(\lambda)$ and the values computed by Eq. (5) is shown in Table II. Values at 0.40, 0.45, etc., are of course correct, so only intermediate values are included in the table. It will be noted that the maximum error is only 1/50 of one percent.

TABLE II. Planckian radiation at 2842°K. $J(\lambda)$ in watt m^{-2} $micron^{-1}$ for Planckian radiator. True values calculated by Planck's equation ($C_2=14320$), approximate values by seven-term polynomial, Eq. (5).

λ	$J(\lambda)$ (true)	$J(\lambda)$ (approx.)	Percent difference
0.425μ	18.924×10^4	18.923×10^4	−0.005
0.475	37.799	37.791	−0.021
0.525	62.940	62.930	−0.016
0.575	92.018	92.003	−0.016
0.625	122.274	122.255	−0.016
0.675	151.245	151.219	−0.017

[1] The international temperature scale is used. See G. K. Burgess, Bur. Stands. J. Res. **1**, 635 (1928); H. T. Wensel, Bur. Stands. J. Res. **22**, 375 (1939); Parry Moon, J. Math. and Phys. **16**, 133 (1937); M. I. T., Elec. Eng. Dept., Contribution No. 131 (1938).

[2] Parry Moon and D. E. Spencer, "Polynomial representation of reflectance curves," J. Opt. Soc. Am. **35**, 597 (1945). See particularly Table V.

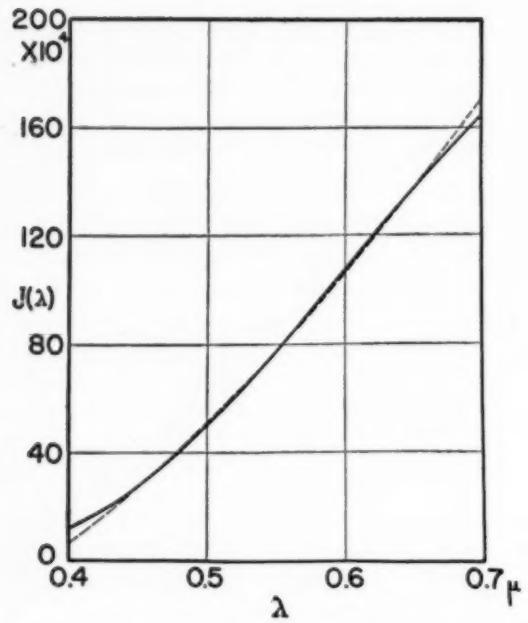


FIG. 2. Planckian radiation at 2842°K, showing the degree of approximation afforded by a three-term polynomial at a common operating temperature for incandescent lamps. ——— Planckian radiation; – – – – three-term analytic approximation.

3. WIEN APPROXIMATION

Another method of approximating the Planckian distribution is by means of the Wien radiation equation. This approximation, while not as simple as the three-term polynomial, allows comparatively easy integration. Since the Planck curve and the Wien curve are of different shapes and cannot be made to coincide, it is necessary to decide what aspect of the two will be made identical. Estey[3] has used color temperature as the criterion by plotting the Wien points in the chromaticity diagram and obtaining shortest distances by Judd's method.[4] Gage[5] used the simple and definite criterion of making the two curves pass through the same points at 0.45 and 0.65μ.

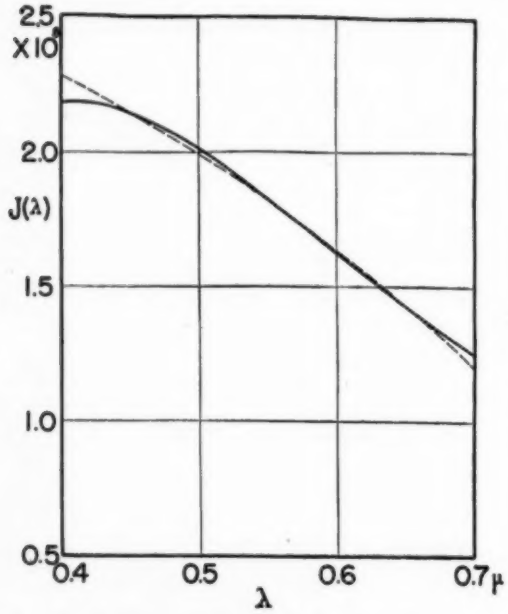


FIG. 3. Planckian radiation at 7000°K. Three-term polynomial approximation is shown dotted.

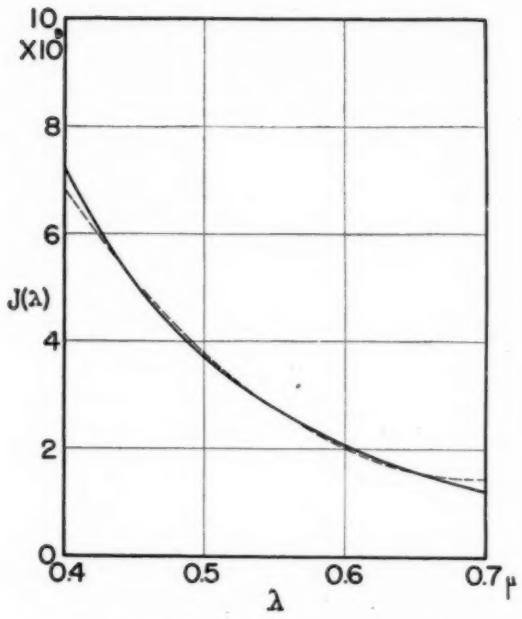


FIG. 4. Planckian radiation at 20,000°K. Three-term polynomial approximation is shown dotted.

We have chosen Gage's method but have altered both T and C_1 in order to obtain coincidence at 0.45 and 0.65μ. Previous work of this kind has been confined entirely to the change in quality, in accordance with the astonishing propensity of colorimetrists to ignore magnitude.

Planck's distribution is expressed in Eq. (2). The Wien distribution is represented by the equation

$$J(\lambda)=\frac{C_1'}{\lambda^5}\exp\left(-C_2/\lambda T'\right), \qquad (6)$$

[3] R. S. Estey, J. Opt. Soc. Am. **28**, 293 (1938).
[4] D. B. Judd, J. Opt. Soc. Am. **26**, 421 (1936).
[5] H. P. Gage, J. Opt. Soc. Am. **23**, 46 (1933); H. P. Gage and N. Macbeth, Trans. I.E.S. **31**, 995 (1936).

where C_2 has the same value as in Eq. (2) but C_1' and T' differ from the Planckian values. If the ordinates of the two curves are to be identical at wave-lengths λ_1 and λ_2 ($\lambda_2 > \lambda_1$),

$$\left.\begin{aligned}\frac{1}{C_1'}\exp(C_2/\lambda_1 T') &= \frac{1}{C_1}[\exp(C_2/\lambda_1 T)-1],\\ \frac{1}{C_1'}\exp(C_2/\lambda_2 T') &= \frac{1}{C_1}[\exp(C_2/\lambda_2 T)-1].\end{aligned}\right\} \quad (7)$$

Dividing, and substituting

$$1/T' = 1/T + 1/\delta, \quad (8)$$

we obtain

$$\frac{1}{\delta} = \frac{\lambda_1\lambda_2}{C_2(\lambda_2-\lambda_1)} \ln\left[\frac{1-\exp(-C_2/\lambda_1 T)}{1-\exp(-C_2/\lambda_2 T)}\right]. \quad (9)$$

The quantity $1/\delta$ is the correction in reciprocal temperature that must be applied in going from the Planckian temperature to the corresponding temperature for the Wien distribution. This correction is always positive, corresponding to universal decrease in temperature from Planck to Wien. The advantage of Eq. (9) over those used by Gage and Estey is that the former is a correction term which is usually small and which therefore needs little accuracy in its computation.

TABLE III. Temperatures for Wien approximation.

T (Planck)	$1/\delta$ (mireds)	ΔT (°K)	T' (Wien)	ΔC_1	C_1' (Wien)
2000°K	0.00167	−0.0067	1999.9933	0.002×10^7	3.6972×10^8
2100	0.00281	0.0124	2099.9876	0.003	3.6973
2200	0.00442	0.0214	2199.9786	0.005	3.6975
2300	0.00697	0.0369	2299.9631	0.008	3.6978
2400	0.01035	0.0596	2399.9404	0.012	3.6982
2500	0.01490	0.0931	2499.9069	0.018	3.6988
2600	0.02085	0.1409	2599.8591	0.025	3.6995
2700	0.02844	0.2073	2699.7927	0.033	3.7003
2800	0.03791	0.2972	2799.7028	0.045	3.7015
2842	0.04252	0.3434	2841.6566	0.050	3.7020
2900	0.04953	0.4165	2899.5835	0.059	3.7029
3000	0.06354	0.5718	2999.4282	0.076	3.7046
3100	0.08018	0.7704	3099.2296	0.096	3.7066
3200	0.09967	1.0203	3198.9797	0.119	3.7089
3300	0.12222	1.3305	3298.6695	0.146	3.7116
3400	0.14803	1.7103	3398.2897	0.177	3.7147
3500	0.17725	2.1700	3497.8300	0.213	3.7183
4000	0.37919	6.0580	3993.9420	0.462	3.7432
4500	0.67991	13.7262	4486.2738	0.841	3.7811
5000	1.0778	26.8020	4973.1980	1.356	3.8326
5500	1.5631	46.8832	5453.1168	2.005	3.8975
6000	2.1216	75.4192	5924.5808	2.780	3.9750
6500	2.7377	113.6445	6386.3555	3.668	4.0638
7000	3.3962	162.5522	6837.4478	4.661	4.1631
20000	17.9979	5293.6610	14706.3391	45.350	8.2320

TABLE IV. Computed values of $J(\lambda)$ (watt m^{-2} micron^{-1}).

λ (micron)	Planck	Wien	Percent difference	Polynomial	Percent difference
	$T=2842.0$°K, $C_1=3.6970\times10^8$,		$T'=2841.66$°K $C_1'=3.7020\times10^8$		
0.40	12.214×10^4	12.212×10^4	−0.017	6.302×10^4	−48.3
0.45	27.475	27.475	0	27.475	0
0.50	49.711	49.716	+0.010	50.866	+ 2.3
0.55	77.159	77.170	+0.014	77.159	0
0.60	107.17	107.18	+0.010	105.65	− 1.4
0.65	137.05	137.05	0	137.05	0
0.70	164.65	164.61	−0.024	170.65	+ 3.6
	$T=7000.0$°K, $C_1=3.6970\times10^8$,		$T'=6837.46$°K $C_1'=4.16312\times10^8$		
0.40	2.1830×10^8	2.1638×10^8	−0.88	2.2830×10^8	+ 4.6
0.45	2.1483	2.1483	0	2.1483	0
0.50	2.0111	2.0204	+0.46	1.9959	− 0.76
0.55	1.8254	1.8359	+0.57	1.8254	0
0.60	1.6254	1.6320	+0.41	1.6370	+ 0.71
0.65	1.4306	1.4306	0	1.4306	0
0.70	1.2508	1.2432	−0.61	1.2065	− 3.5
	$T=20{,}000$°K, $C_1=3.6970\times10^8$,		$T'=14{,}706.3$°K $C_1'=8.23200\times10^8$		
0.40	7.2360×10^9	7.0468×10^9	−2.6	6.7731×10^9	− 6.4
0.45	5.1251	5.1251	0	5.1251	0
0.50	3.7120	3.7574	+1.2	3.7827	+ 1.9
0.55	2.7451	2.7847	+1.4	2.7451	0
0.60	2.0689	2.0890	+0.97	2.0134	− 2.7
0.65	1.5862	1.5862	0	1.5862	0
0.70	1.2350	1.2187	−1.3	1.4651	+18.6

This advantage is particularly noticeable at the lower temperatures.

Equation (9) is exact. Usually, however, the exponentials are small in comparison with unity and a simpler expression is possible. Since

$$\ln(1-x) = -[x+\tfrac{1}{2}x^2+\tfrac{1}{3}x^3+\cdots],$$

Eq. (9) may be written,

$$\frac{1}{\delta} \cong \frac{\lambda_1\lambda_2}{C_2(\lambda_2-\lambda_1)} \times[\exp(-C_2/\lambda_2 T)-\exp(-C_2/\lambda_1 T)]. \quad (10)$$

If $\lambda_1=0.450\mu$, $\lambda_2=0.650\mu$, and $C_2=14{,}320$, Eq. (9) becomes

$$\frac{1}{\delta} = 1.0213\times10^{-4} \ln\left[\frac{1-\exp(-31{,}822/T)}{1-\exp(-22{,}031/T)}\right] \quad (9a)$$

and Eq. (10) is

$$\frac{1}{\delta} \cong 1.0213\times10^{-4}[\exp(-22{,}031/T) - \exp(-31{,}822/T)]. \quad (10a)$$

Next the constant C_1' is evaluated. From Eqs. (2) and (6),

$$C_1' = C_1 \frac{\exp(C_2/\lambda_1 T')}{[\exp(C_2/\lambda_1 T)-1]}. \quad (11)$$

Let

$$C_1' = C_1 + \Delta C_1.$$

Then the increment in C_1, to be used in changing from Planck to Wien radiation, is

$$\Delta C_1 = C_1' - C_1$$

or

$$\Delta C_1 = C_1 \times \left[\frac{\exp(C_2/\lambda_1\delta) - 1 + \exp(-C_2/\lambda_1 T)}{1 - \exp(-C_2/\lambda_1 T)}\right]. \quad (12)$$

Equation (12) is exact. At temperatures below approximately 3500°K it may be approximated by

$$\Delta C_1 \leqq C_1\left[\frac{C_2}{\lambda_1}\frac{1}{\delta} + \exp(-C_2/\lambda_1 T)\right]. \quad (12a)$$

Below 2800°K, the exponential may be neglected, resulting in the simple equation,

$$\Delta C_1 \leqq \frac{C_1 C_2}{\lambda_1}\frac{1}{\delta}. \quad (12b)$$

If $\lambda_1 = 0.45\mu$, $C_1 = 3.6970 \times 10^8$, $C_2 = 14{,}320$, Eq. (12) becomes

$$\Delta C_1 = 3.6970 \times 10^8 \times \left[\frac{\exp(31822 \cdot 1/\delta) - 1 + \exp(-31822/T)}{1 - \exp(-31822/T)}\right] \quad (13)$$

and Eq. (12a) reduces to

$$\Delta C_1 \leqq 3.6970 \times 10^8 [31822 \cdot (1/\delta) + \exp(-31822/T)]. \quad (13a)$$

Also Eq. (12b) becomes

$$\Delta C_1 \leqq 11.765 \times 10^{12}(1/\delta). \quad (13b)$$

Values of $1/\delta$ and ΔC_1 are listed in Table III. Also given are values of T', $\Delta T'$, and C_1'. A comparison of computed values of $J(\lambda)$ is shown in Table IV. The values calculated by Planck's equation are listed in the second column, and the Wien values are in the third column. Computations made by use of the three-term polynomial approximation are also given (fifth column). Table IV shows that at 2842°K, the maximum error caused by using the Wien equation instead of the Planck equation is only 1/40 percent.

TABLE V. Radiant pharosage in visible spectrum. D_r (watt m^{-2}) for Planckian radiator.

	$T = 2842$°K	7000	20,000
Planck	2.4290×10^5		
Wien	2.4290×10^5	5.3933×10^7	9.6566×10^8
2-term polynomial	2.4264×10^5	5.3929×10^7	
(2-term-Wien)/Wien	0.0011	0.00007	
3-term polynomial	2.4244×10^5	5.3955×10^7	9.6090×10^8
(3-term-Wien)/Wien	0.0019	0.00041	0.0049
D_r(visible region)/D_r(total)	0.0652	0.3935	0.0106

For higher temperatures, the discrepancies between the Wien and Planck distributions are more pronounced. For 20,000°K, differences of one or two percent are encountered at some wave-lengths between the true curve and the correlated Wien curve ($T' = 14{,}706$, $C_1' = 8.232 \times 10^8$). The Wien distribution is always too peaked, giving values that are too high in the center and too low at the ends of the visible spectrum. The three-term polynomial representation is inferior in accuracy to the Wien distribution at low temperatures; but at temperatures above 7000°K, the two are not widely different.

4. INTEGRATION WITHOUT WEIGHTING FUNCTION

It is sometimes necessary to obtain the radiant power in a specified wave-length band. For instance, the portion of the radiant power of a Planckian radiator may be required for the visible region or for a portion of the visible region between the wave-lengths λ_1 and λ_2. The Planck equation then proves to be inconvenient since it can be integrated only by expansion in an infinite series of exponentials. A numerical method of integration was developed by Holladay,[6] based on the tables of Frehafer and Snow.[7] The work differs completely from the present paper in that analytic methods were not used.

The integral to be evaluated is

$$D_r = \int_{\lambda_1}^{\lambda_2} J(\lambda)\, d\lambda, \quad (14)$$

where D_r = radiant pharosage (watt m^{-2}) between the wave-lengths λ_1 and λ_2, and $\lambda_2 > \lambda_1$ (micron).

[6] L. L. Holladay, J. Opt. Soc. Am. **17**, 329 (1928).
[7] M. K. Frehafer and C. S. Snow, Miscellaneous Publications of the Bureau of Standards, No. 56 (1925).

If a three-term polynomial approximation is employed,

$$J(\lambda)=K_0+K_1\lambda+K_2\lambda^2, \tag{3}$$

and the integration is very simple:

$$D_r=\int_{\lambda_1}^{\lambda_2}(K_0+K_1\lambda+K_2\lambda^2)d\lambda,$$

$$=K_0(\lambda_2-\lambda_1)+\frac{K_1}{2}(\lambda_2^2-\lambda_1^2)+\frac{K_2}{3}(\lambda_2^3-\lambda_1^3). \tag{15}$$

As an example of the use of Eq. (15), Table V was computed. It gives the radiant pharosage in the wave-length band from 0.40 to 0.70μ for Planckian radiators operating at several temperatures. The values calculated by means of Eq. (15) are compared with the correct values obtained by series integration of Planck's equation. The discrepancies are seen to be less than a half percent in all cases.

Now consider the integration of the generalized Wien equation:

$$J(\lambda)=C_1\lambda^{-5}\exp(-nC_2/\lambda T). \tag{16}$$

Equation (14) becomes

$$D_r=\int_{\lambda_1}^{\lambda_2}J(\lambda)d\lambda=C_1\int_{\lambda_1}^{\lambda_2}\lambda^{-5}\exp(-nC_2/\lambda T)d\lambda$$

or

$$D_r=C_1\left(\frac{T}{nC_2}\right)^4\left\{\exp(-nC_2/\lambda_2T)\cdot\left[\left(\frac{nC_2}{\lambda_2T}\right)^3+3\left(\frac{nC_2}{\lambda_2T}\right)^2+6\left(\frac{nC_2}{\lambda_2T}\right)+6\right]\right.$$

$$\left.-\exp(-nC_2/\lambda_2T)\cdot\left[\left(\frac{nC_2}{\lambda_1T}\right)^3+3\left(\frac{nC_2}{\lambda_1T}\right)^2+6\left(\frac{nC_2}{\lambda_1T}\right)+6\right]\right\}. \tag{17}$$

For Planckian radiation, a series expansion is employed:

$$J(\lambda)=\frac{C_1}{\lambda^5}\frac{1}{\exp(C_2/\lambda T)-1},$$

$$=\frac{C_1}{\lambda^5}[\exp(-C_2/\lambda T)+\exp(-2C_2/\lambda T)+\cdots], \tag{18}$$

and the integration of each term is obtained by use of Eq. (17) which becomes

$$D_r=0.87918(T/n)^4\left\{\exp(-14320n/\lambda_2T)\cdot\left[\left(\frac{14320n}{\lambda_2T}\right)^3+3\left(\frac{14320n}{\lambda_2T}\right)^2+6\left(\frac{14320n}{\lambda_2T}\right)+6\right]\right.$$

$$\left.-\exp(-14320n/\lambda_1T)\cdot\left[\left(\frac{14320n}{\lambda_1T}\right)^3+3\left(\frac{14320n}{\lambda_1T}\right)^2+6\left(\frac{14320n}{\lambda_1T}\right)+6\right]\right\}. \tag{17a}$$

Evidently Eq. (18) leads to great complication, particularly at high temperatures. Simplification is effected by using the Wien equation ($n=1$) with temperature T' and constant C_1'. Then Eq. (17) becomes

$$D_r=C_1'(T'/14320)^4\left\{\exp(-14320/\lambda_2T)\cdot\left[\left(\frac{14320}{\lambda_2T}\right)^3+3\left(\frac{14320}{\lambda_2T}\right)^2+6\left(\frac{14320}{\lambda_2T}\right)+6\right]\right.$$

$$\left.-\exp(-14320/\lambda_1T)\cdot\left[\left(\frac{14320}{\lambda_1T}\right)^3+3\left(\frac{14320}{\lambda_1T}\right)^2+6\left(\frac{14320}{\lambda_1T}\right)+6\right]\right\}. \tag{17b}$$

TABLE VI. Coefficients for integration of polynomial approximation, light sources.

m	X/K_m	Y/K_m	Z/K_m
0	^{0}10666 57240	^{0}10685 35175	^{0}10677 25053
1	^{1}60921 26156	^{1}59773 82336	^{1}48473 18484
2	^{1}35226 72563	^{1}33616 34585	^{1}22066 86373
3	^{1}20585 29574	^{1}19006 76935	^{1}10073 48896
4	^{1}12138 18444	^{1}10804 00986	^{2}46112 93726
5	^{2}72128 04822	^{2}61741 99087	^{2}21167 64643
6	^{2}43147 45424	^{2}35472 77575	^{2}97438 95977
7	^{2}25962 45164	^{2}20489 36516	^{2}44978 64475
8	^{2}15703 36300	^{2}11898 18699	^{3}20820 78752
9	^{3}95428 23493	^{3}69462 88390	^{4}96651 46038
10	^{3}67101 94957	^{3}40770 52600	^{4}44992 86102
0	^{0}10666 57240	^{0}10685 35175	^{0}10677 25053
− 1	^{0}18945 52333	^{0}19203 57729	^{0}23583 64653
− 2	^{0}34207 93241	^{0}34696 72926	^{0}52233 94835
− 3	^{0}62918 49167	^{0}63023 83522	1.1600 63798
− 4	1.1809 92005	1.1508 72075	2.5834 18571
− 5	2.2652 57902	2.1127 61624	5.7688 38040
− 6	4.4431 71306	3.8991 48094	12.916 93227
− 7	8.9117 26706	7.2340 01869	29.000 44752
− 8	18.261 13996	13.491 81539	65.286 16073
− 9	38.177 40910	25.295 17554	147.36 87925
−10	81.275 83667	47.673 01938	333.54 52442

Note.—Superscripts refer to the number of zeros after the decimal point. Thus $^{2}7212804822 = 0.007212804822$.

Values of D_r computed by means of Eq. (17b) for $\lambda_1 = 0.40$ and $\lambda_2 = 0.70\mu$ are listed in Table V. At 2842°K, the value obtained from Eq. (17b) agrees with the true Planckian value, Eqs. (16a) and (17), in the fifth digit. The Planckian results are

$$n=1, \quad D_r = 2.4283\times10^5;$$
$$n=2, \quad D_r = 0.0007\times10^5;$$
$$n=3, \quad D_r = \text{negligible}.$$

The summation of these values gives the true result, $D_r = 2.4290\times10^5$ watt m^{-2}. Note that the results obtained by using our corresponding T' and C_1' are correct to five figures, while the common practice of assuming that $T'=T$ and $C_1'=C_1$ gives a result that is low by 7 in the fifth figure. At higher temperatures the exact computation becomes very laborious and the approximate methods are correspondingly advantageous.

5. INTEGRATION WITH WEIGHTING FUNCTIONS

Even more important than the integrations described in the preceding section are the weighted integrations that occur so frequently in colorimetry. If the weighting functions are $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, $\bar{z}(\lambda)$, then

$$X = \int_0^\infty \bar{x}(\lambda)J(\lambda)d\lambda,$$
$$Y = \int_0^\infty \bar{y}(\lambda)J(\lambda)d\lambda, \qquad (19)$$
$$Z = \int_0^\infty \bar{z}(\lambda)J(\lambda)d\lambda.$$

If $J(\lambda)$ is expressed as a polynomial,

$$J(\lambda) = \sum_m K_m\lambda^m \qquad (20)$$

and the weighting functions are given by their analytic representations,[8] then the integrals are

TABLE VII. Three-term polynomial approximations.

	T = 2000°K	2842	3500	7000	20,000
X, Planck	2812.4	97,871	^{1}49221	^{3}18413	^{4}28581
Polynomial approximation	2834.5	96,917	^{1}49007	^{3}18490	^{4}28219
Percent difference	+0.79	+0.98	−0.44	+0.42	−1.3
Y, Planck	2203.0	89,030	^{1}47543	^{3}19073	^{4}28727
Polynomial approximation	2201.0	88,721	^{1}47531	^{3}19077	^{4}28705
Percent difference	−0.091	−0.34	−0.025	+0.021	−0.077
Z, Planck	326.03	31,808	^{1}25058	^{3}22738	^{4}54090
Polynomial approximation	343.04	31,371	^{1}24905	^{3}22794	^{4}53830
Percent difference	+4.5	−1.4	−0.61	+0.25	−0.48
x, Planck	0.52653	0.44749	0.40404	0.30574	0.25656
Polynomial	0.52700	0.44660	0.40354	0.30632	0.25479
Difference	$+4.7\times10^{-4}$	-8.9×10^{-4}	-5.0×10^{-4}	$+5.8\times10^{-4}$	-17.7×10^{-4}
y, Planck	0.41243	0.40707	0.39027	0.31671	0.25788
Polynomial	0.40922	0.40884	0.39138	0.31605	0.25918
Difference	-32.1×10^{-4}	$+17.7\times10^{-4}$	$+11.1\times10^{-4}$	-6.6×10^{-4}	$+13.0\times10^{-4}$

Note.—Superscripts denote the number of zeros before the decimal point. For instance, $^{1}49221 = 492{,}210$.

[8] Parry Moon and D. E. Spencer, "Analytical representation of trichromatic data," J. Opt. Soc. Am. **35**, 399 (1945); and "Analytic expressions in photometry and colorimetry," to appear in J. Math. and Phys.

evaluated analytically by use of the formulas,

$$(L/K)_m = A\frac{\Gamma(p-m-1)}{q^{p-m-1}},$$

$$(L/K)_{m+1} = (L/K)_m \frac{q}{p-m-2}, \qquad (21)$$

$$(L/K)_{m-1} = (L/K)_m \frac{p-m-1}{q},$$

where L represents X, Y, or Z and where p and q are the constants given previously.[8] These equations were used in the calculation of the constants of Table VI.

To obtain the trichromatic specification of any radiation whose spectral distribution can be approximated by a polynomial, one multiplies the polynomial coefficients K_m by the corresponding values of Table VI and adds:

$$L = \sum_m (L/K_m)K_m. \qquad (22)$$

For Planckian radiation, the values of K_m for a three-term polynomial approximation may be employed. Examples are given in Table VII. Generally the error in X, Y, or Z caused by the polynomial approximation is less than one percent. The homogeneous coordinates x and y are also given in the table and the discrepancy caused by polynomial approximation is usually less than 20×10^{-4}.

If greater accuracy is desired, a seven-term polynomial may be employed. Table VIII shows that at 2842°K the trichromatic coordinates X, Y, and Z are in error because of polynomial approximation by less than 1/50 percent, while x and y are within 0.1×10^{-4} of the Planckian value obtained by the analytic method.

When $J(\lambda)$ is expressed in terms of a corresponding Wien distribution

$$J(\lambda) = \frac{C_1'}{\lambda^5}\exp(-C_2/\lambda T'), \qquad (23)$$

the trichromatic coefficients of the radiation are found by evaluating the expression[8]

$$L = \frac{AC_1'\Gamma(p+4)}{(q+C_2/T')^{p+4}} \qquad (24)$$

TABLE VIII. Seven-term polynomial approximation at 2842°K.

X, Planck	97,871
Polynomial approximation	97,856
Percent difference	−0.015
Y, Planck	89,030
Polynomial approximation	89,016
Percent difference	−0.017
Z, Planck	31,808
Polynomial approximation	31,802
Percent difference	−0.019
x, Planck	0.44749
Polynomial approximation	0.44750
Difference	$+0.1\times10^{-4}$
y, Planck	0.40707
Polynomial approximation	0.40707
Difference	0.0×10^{-4}

for Z, X_B, X_C, Y_A, Y_D, and X_A and by using the relations

$$X = X_A + X_C - X_B,$$
$$Y = Y_A - Y_D. \qquad (25)$$

The corresponding values of C_1' and T' for Planckian distributions, Table IV, were used in evaluating X, Y, and Z. The results are shown in Table IX. It will be noted that X, Y, and Z are always within one percent of the values obtained with Planckian distributions. As would be expected, the errors are only a few thousandths of one percent at low temperatures. The homogeneous coordinates are accurate to five decimal places at low temperatures, and at 20,000°K are still accurate enough for practical purposes.

The total lamprosity $\bar{y}_l$ (luminous efficacy) is also computed very easily by the methods of this paper. By definition,

$$\bar{y}_l = \int_0^\infty \bar{y}(\lambda)J(\lambda)d\lambda \Big/ \int_0^\infty J(\lambda)d\lambda. \qquad (26)$$

Multiplication[8] by 648 gives lumens per watt. The numerator is obtained from Eq. (21) or (23), while the denominator is obtained from the Stefan-Boltzmann relation,

$$\int_0^\infty J(\lambda)d\lambda = 5.709\times10^{-8}T^4 \text{ watt/m}^2. \qquad (27)$$

When Planckian radiation is transmitted through a selective medium or reflected from a surface, the additional weighting function $\tau(\lambda)$ or

TABLE IX. Comparison of Planck and Wien equations, where Wien temperature and C_1' are adjusted in accordance with Eqs. (9) and (12).

	$T=2000°K$	2842	3500	7000	20,000
X, Planck	2812.4	97,871	49221[1]	18413[3]	28581[4]
Wien	2812.5	97,879	49237[1]	18467[3]	28741[4]
Percent difference	+0.0036	+0.0082	+0.037	+0.29	+0.56
Y, Planck	2203.0	89,030	47543[1]	19073[3]	28727[4]
Wien	2203.0	89,040	47563[1]	19165[3]	29079[4]
Percent difference	0	+0.011	+0.042	+0.48	+0.87
Z, Planck	326.03	31,808	25058[1]	22738[3]	54090[4]
Wien	326.03	31,808	25059[1]	22737[3]	54013[4]
Percent difference	0	0	+0.0040	−0.0044	−0.14
x, Planck	0.52653	0.44749	0.40404	0.30574	0.25656
Wien	0.52653	0.44749	0.40405	0.30590	0.25700
Difference	0.0×10^{-4}	0.0×10^{-4}	$+0.1\times10^{-4}$	$+1.6\times10^{-4}$	$+4.4\times10^{-4}$
y, Planck	0.41243	0.40707	0.39027	0.31671	0.25788
Wien	0.41243	0.40708	0.39031	0.31746	0.26002
Difference	0.0×10^{-4}	$+0.1\times10^{-4}$	$+0.4\times10^{-4}$	-7.5×10^{-4}	$+21.4\times10^{-4}$

$\rho(\lambda)$ is introduced. If the spectral reflectance or transmittance curves are represented also by polynomials, the integrations can be handled as in the preceding paragraphs. This subject is treated in a previous publication.[2] The total luminous reflectance is

$$\rho_l=\int_0^\infty \rho(\lambda)\bar{y}(\lambda)J(\lambda)d\lambda \Big/ \int_0^\infty \bar{y}(\lambda)J(\lambda)d\lambda, \quad (28)$$

where numerator and denominator are evaluated analytically. A similar expression applies to the total luminous transmittance τ_l.

6. CONCLUSIONS

An investigation has been made of the possibility of simplifying computations dealing with Planckian radiation. It is found that Planckian distributions can be approximated in the visible region by three-term polynomials. Integrations can be made analytically with surprising ease and the results are generally correct to within one percent. Values of the coordinates x and y in the chromaticity diagram obtained by this method are in error ordinarily by less than 20×10^{-4}.

If greater accuracy is desired, it can be obtained by the same methods, using a larger number of terms in the polynomial. Data are given for a seven-term approximation which gives accuracy to five figures.

The Planckian distribution can be approximated also by a Wien distribution. The temperature is altered and a modified constant C_1 is used. Equations and tables are given so that the corresponding temperatures and values of C_1 can be obtained. The Wien distribution is more accurate than the three-term polynomial but is more trouble to compute.

Calendar of Meetings

June

20–22 American Association of Physics Teachers, St. Louis, Missouri
20–22 American Physical Society, Chicago, Illinois
20–23 Society for the Promotion of Engineering Education, St. Louis, Missouri
24–28 American Institute of Electrical Engineers, Detroit, Michigan
24–28 American Society for Testing Materials, Buffalo, New York

July

12–13 American Physical Society, Berkeley, California
17–19 American Society of Civil Engineers, Spokane, Washington

August

14–15 Institute of Aeronautical Sciences, Los Angeles, California
19–20 Mathematical Association of America, Ithaca, New York
20–23 American Mathematical Society, Ithaca, New York

September

9–13 American Chemical Society, Chicago, Illinois
16–20 The Instrument Society of America, Pittsburgh, Pennsylvania
18–21 Illuminating Engineering Society, Quebec, Canada
19–21 American Physical Society, New York, New York
30–October 3 American Society of Mechanical Engineers, Boston, Massachusetts

October

3– 5 National Electronics Conference, Inc., Chicago, Illinois
9–11 American Society of Civil Engineers, Kansas City, Missouri
25 American Mathematical Society, New York, New York

Approximate Calculation of the Range of a Winged Jet Projectile

M. Zbigmew Krzywoblocki
Palo Alto, California

(Received January 2, 1946)

The purpose of this note is to discuss the possibility of launching winged jet bombs or mail-rockets on very long flights of perhaps 3000 miles, but the problem is considered from one standpoint only, i.e., that of mass ratio. There will be no discussion of the directional and stability instruments of a long distance winged jet projectile and the real mechanics of flight of such a unit, nor on various problems connected with winds, accuracy, etc. The problem will be treated in simplified form from the point of view of mass ratio at the beginning and at the end of flight. The simplified equations of 3 periods of flight are given: climbing, horizontal flight, and gliding. The simplifications introduced present a very rough approximation to the actual problem, but exact solution is not possible because of mathematical difficulties. A short example at the end of this note shows that the required mass ratio for a flight of 2000 or 3000 miles is high.

1. CLIMBING

ASSUME the flight path of a winged jet projectile consuming external air to be a straight line inclined to the level at an angle ϕ. The initial velocity is equal to zero. The reactional force of the exhaust gases is

$$F=m_f[c+\alpha(c-v)], \tag{1}$$

where

F, reactional force of the outflowing gases.
c, velocity of the outflowing gases with respect to the nozzle.
m_f, mass of the propulsive material (fuel).
α, air/fuel ratio by weight or by mass.
v, velocity of flight.

Substitute the following expression

$$m_f=(1/g)kW, \tag{2}$$

where

W, instantaneous weight of the jet bomb.
k, coefficient.
g, acceleration of gravity at sea level.

Neglect as a first approximation the influence of height on the value of the acceleration of gravity. Thus

$$F=(1/g)kW[c+\alpha(c-v)]. \tag{3}$$

The rate of fuel injection is

$$dW=-kWdt. \tag{4}$$

From (4) with the initial condition, i.e., for $t=0$, $W=W_0$, one obtains

$$\frac{W}{W_0}=\exp\left[-\int_0^t kdt\right]. \tag{5}$$

Express the change in density of the air by the Hohmann's formula[1]

$$\frac{\rho}{\rho_0}=\left[1-\frac{h}{400{,}000}\right]^{49}, \tag{6}$$

where

ρ, density at the given height.
ρ_0, density at sea level.
h, height.

The lift at any height may be given by the formula

$$L=W\cos\phi=(\rho/2)c_LSv^2, \tag{7}$$

where

L, lift.
c_L, lift coefficient.
S, wing area.

At sea level one has

$$W_0=(\rho_0/2)c_LSv_0^2, \tag{8}$$

where v_0 denotes the velocity of horizontal flight at sea level. In a motion with subsonic velocity it is permissible to assume that the coefficient c_L is constant. Thus one obtains:

$$\left(\frac{W}{W_0}\right)\cos\phi=\frac{L}{W_0}=\left(\frac{v^2}{v_0^2}\right)\left[1-\frac{h}{400{,}000}\right]^{49}. \tag{9}$$

Similarly the drag may be expressed in the form

$$D=\frac{\rho}{2}c_DSv^2=\epsilon L, \tag{10}$$

[1] Eugen Sänger, *Raketen flugtechnik* (Verlag von R. Oldenbourg, Müchen and Berlin, 1933).

where

c_D, drag coefficient assumed to be constant in subsonic flight.
ϵ, c_D/c_L.

From the equation of equilibrium of forces in the direction of motion one obtains

$$\frac{F}{W}=\sin\phi+\epsilon\cos\phi+\left(\frac{1}{g}\right)\left(\frac{dv}{dt}\right). \quad (11)$$

Assume as a first, very rough approximation that W is a constant and equal to the mean of the values of the weight of the rocket at the beginning and end of the climbing period[1]

$$W=\frac{W_0+W_1}{2}, \quad \frac{W}{W_0}=\frac{W_0+W_1}{2W_0}=k_1=\text{const.}, \quad (12)$$

where

W_1, weight of the jet projectile at the end of the climbing period.

From (9) one obtains

$$v=\frac{ds}{dt}=v_0(k_1\cos\phi)^{\frac{1}{2}}\left[1-\frac{s\sin\phi}{400{,}000}\right]^{-24.5}, \quad (13)$$

where

s, distance of flight measured along the path.
t, time.
h, $s\sin\phi$.

Integration of (13) between the limits 0 and t with the initial condition $t=0$ for $s=0$ gives

$$t=\left[\frac{15{,}689}{v_0(k_1\cos\phi)^{\frac{1}{2}}\sin\phi}\right]\times\left[1-\left(1-\frac{s\sin\phi}{400{,}000}\right)^{25.5}\right], \quad (14)$$

$$s=\frac{400{,}000}{\sin\phi}\times\left[1-\left(1-\frac{v_0t\sin\phi(k_1\cos\phi)^{\frac{1}{2}}}{15{,}689}\right)^{1/25.5}\right], \quad (15)$$

$$\frac{dv}{dt}=\frac{dv}{ds}\frac{ds}{dt}=\frac{dv}{ds}v=\left[\frac{v_0^2k_1\sin 2\phi}{32{,}640}\right]\left[1-\frac{s\sin\phi}{400{,}000}\right]^{-50}. \quad (16)$$

Substitute (16) into (11) and compare (11) with (3)

$$\left(\frac{1}{g}\right)k[c+\alpha(c-v)]=\sin\phi+\epsilon\cos\phi+\left[\frac{v_0^2k_1\sin 2\phi}{32{,}640g}\right]\left[1-\frac{s\sin\phi}{400{,}000}\right]^{-50}. \quad (17)$$

Substitute (15) into (17) and obtain

$$k=\left[\frac{g}{c(\alpha+1)-\alpha v}\right][\sin\phi+\epsilon\cos\phi+AB^{-1.96}], \quad (18)$$

$$A=\frac{v_0^2k_1\sin 2\phi}{32{,}640g}=\frac{E}{g}, \quad (18a)$$

$$B=1-\frac{v_0t\sin\phi(k_1\cos\phi)^{\frac{1}{2}}}{15{,}689}. \quad (18b)$$

Substitute (13) into (14) and obtain

$$t=\left[\frac{15{,}689}{v_0\sin\phi(k_1\cos\phi)^{\frac{1}{2}}}\right]\times\left[1-\left(\frac{v_0(k_1\cos\phi)^{\frac{1}{2}}}{v}\right)^{1.04}\right]. \quad (19)$$

By the use of (5), the final expression for the mass ratio at the end of climbing is

$$\frac{W_1}{W_0}=2k_1-1=\exp\left\{-\int_0^{t_1}\frac{g[\sin\phi+\epsilon\cos\phi+AB^{-1.96}]}{c(\alpha+1)-CB^{-0.96}}dt\right\}, \quad (20)$$

$$C=av_0(k_1\cos\phi)^{\frac{1}{2}}=\alpha H; \quad (20a)$$

or

$$\frac{W_1}{W_0}=\exp\left\{-\int_0^{v_1}\frac{g[\sin\phi+\epsilon\cos\phi+A(v/H)^{2.04}]G^{50}}{[c(\alpha+1)-\alpha v]E}dv\right\}, \quad (21)$$

$$G=\left(1-\frac{s_1 \sin \phi}{400{,}000}\right), \tag{21a}$$

where v_1 is the velocity at the end of climbing:

$$v_1=HG^{-24.5}, \tag{22}$$

$$l_1=s_1 \cos \phi=h \cot \phi, \tag{23}$$

where the symbols denote

l_1, horizontal distance traversed during climbing.
s_1, total length of flight path during climbing.

2. HORIZONTAL FLIGHT

To obtain the equations of horizontal flight, put in the above formulas the following values:

$$h=\text{const.}, \quad \phi=0, \quad v=v_h.$$

As the initial values for this flight assume

$$W=W_1, \quad s=s_1, \quad v=v_1';$$

and as the end values:

$$t=t_2, \quad W=W_2, \quad s=s_2, \quad v=v_2.$$

Assume $\epsilon=\text{const.}$

$$\frac{W}{W_1}=\left(\frac{v}{v_1'}\right)^2, \quad \frac{W}{W_1}=\frac{W_1+W_2}{2W_1}=k_2=\text{const.}, \tag{24}$$

$$W_1=\frac{\rho_h}{2}c_L S v_1'^2,$$

$$v=\frac{ds}{dt}=v_1'(k_2)^{\frac{1}{2}}, \quad (t-t_1)=\left[\frac{s-s_1}{v_1'(k_2)^{\frac{1}{2}}}\right], \tag{25}$$

$$(s-s_1)=v_1'(k_2)^{\frac{1}{2}}(t-t_1), \quad \frac{dv}{dt}=0, \tag{26}$$

or

$$k=\left[\frac{g\epsilon}{c(\alpha+1)-\alpha v}\right], \tag{27}$$

$$\frac{W}{W_1}=\exp\left(-\int_{t_1}^{t} k dt\right)=\exp\left[-\frac{g\epsilon(t-t_1)}{c(\alpha+1)-\alpha v}\right]$$

$$=\exp\left[-\frac{g\epsilon(t-t_1)}{c(\alpha+1)-\alpha v_1'(k_2)^{\frac{1}{2}}}\right]. \tag{28}$$

At the end of horizontal flight one has

$$\frac{W_2}{W_1}=2k_2-1=\exp\left[-\frac{g\epsilon(t_2-t_1)}{c(\alpha+1)-\alpha v_1'(k_2)^{\frac{1}{2}}}\right]$$

$$=\exp\left[-\frac{g\epsilon(s_2-s_1)}{v_1'(k_2)^{\frac{1}{2}}c(\alpha+1)-\alpha v_1'^2 k_2}\right], \tag{29}$$

$$(s_2-s_1)=v_1'(k_2)^{\frac{1}{2}}(t_2-t_1). \tag{30}$$

3. GLIDING

During gliding there is no reactional force, and the weight of the projectile remains unchanged. In the case where the flight path is a curved line and the velocity of flight is very great, one has to take into account the centrifugal force. In flight with supersonic velocity the lift coefficient is a complicated function of several parameters, but as Sänger showed it may be given in a first very rough approximation by the formula

$$c_{LS}=\left(\frac{165{,}300}{v^2}\right)+0.01. \tag{31}$$

Assume as the initial values $t=t_2$, $v=v_2$, $s=s_2$; and as the end values $t=t_3$, $v=v_3$, $s=s_3$. Apply the following symbols:

$$F_c=Wv^2/(Rg), \tag{32}$$

$$L=\frac{\rho_0}{2}S\left(\frac{165{,}300}{v^2}+0.01\right)\left(1-\frac{h}{400{,}000}\right)^{49}v^2, \tag{33}$$

$$D=L\epsilon, \quad W=\text{const.}, \quad T=(W/g)(dv/dt), \tag{34}$$

where

F_c, centrifugal force;
T, inertia force in the tangential direction;
R, average radius of the earth.

Use the substitution:

$$p=\rho_0 S/(2W)=\text{const.}, \quad \epsilon=\text{const.}, \tag{35}$$

$$L=pW\left(\frac{165{,}300}{v^2}+0.01\right)\left(1-\frac{h}{400{,}000}\right)^{49}v^2. \tag{36}$$

The equilibrium of forces in the radial and tangential directions gives two conditions:

$$p\left(\frac{165{,}300}{v^2}+0.01\right)\left(1-\frac{h}{400{,}000}\right)^{49}v^2$$

$$+\left(\frac{v^2}{Rg}\right)=\cos\phi, \tag{37}$$

$$\epsilon\rho\left(\frac{165{,}300}{v^2}+0.01\right)\left(1-\frac{h}{400{,}000}\right)^{49}v^2 = \sin\phi+\frac{1}{g}\left(\frac{dv}{dt}\right), \quad (38)$$

where ϕ is the instantaneous angle between the tangent to the path at an arbitrary point and the horizontal. It is assumed that the motion is decelerated.

Eliminating h from these two equations

$$\frac{d^2s}{dt^2}=\frac{dv}{dt}=g\epsilon\cos\phi-g\sin\phi-\frac{v^2}{R}\epsilon. \quad (39)$$

In order to solve this equation in an elementary way assume for ϕ a constant mean value. The difference between the straight flight path and the curved one is that the term representing the centrifugal force remains. After the first integration one obtains

$$t-t_2=\left(\frac{1}{2K}\right)\ln\left[N\left(\frac{M+vK}{M-vK}\right)\right], \quad (40)$$

$$K=\left(\frac{\epsilon^2}{R}g\cos\phi-\frac{\epsilon}{R}g\sin\phi\right)^{\frac{1}{2}}, \quad (40a)$$

$$N=(M-v_2K)/(M+v_2K), \quad (40b)$$

$$M=g\epsilon\cos\phi-g\sin\phi; \quad (40c)$$

or

$$v=\left(\frac{M}{K}\right)\frac{\{\exp[2(t-t_2)K]-N\}}{\{\exp[2(t-t_2)K]+N\}}. \quad (41)$$

The second integration gives

$$(s-s_2)=\left(\frac{M}{K^2}\right)\ln\left\{\frac{N+\exp[2(t-t_2)K]}{N+1}\right\}-\left(\frac{M}{K}\right)(t-t_2). \quad (42)$$

The horizontal distance covered during this period may be calculated as a first approximation from the value $(s-s_2)$ and from the chosen value of the angle ϕ.

If the flight path is assumed to be a straight line one obtains from (39)

$$v-v_2=g(\epsilon\cos\phi-\sin\phi)(t-t_2), \quad (43)$$

$$s-s_2=(g/2)(\epsilon\cos\phi-\sin\phi)(t-t_2)^2+v_2(t-t_2). \quad (44)$$

The horizontal distance during gliding is

$$l_3=(s-s_2)\cos\phi. \quad (45)$$

In case the velocity of flight is smaller than the velocity of sound the coefficient c_L is constant. Another way to approximately calculate the horizontal distance is

$$l_3=K_2/D=K_2/W_2, \quad (46)$$

where K_2 is the kinetic energy at the end of horizontal flight. Assuming for D or ϵ a mean value during gliding one may obtain by short calculation the desired result.

Example

For the sake of simplicity only the distance covered in the horizontal flight period will be calculated. Assume the coefficient of internal efficiency of the motor to be 0.25, which is fairly high. As is known this coefficient is given by the expression

$$\eta_i=c^2/c_t^2, \quad (47)$$

where

c, real velocity of the exhaust gases.
c_t, theoretical velocity of the exhaust gases.

Consequently in the formulas employed, the real velocity c i.e., $c=0.5c_t$, will be included instead of the theoretical velocity.

Use formula (29) with

$$\alpha=15, \quad \epsilon=\tfrac{1}{5},^* \quad c_t=10{,}000 \text{ ft./sec.}, \quad v_1'=300 \text{ ft./sec.}$$

---

* The tests of several airfoils at high speeds showed the following limits of ϵ at $V/c_s=1.08$ (c_s is the speed of sound at temperature of jet) and at 0° or 2° of angle of attack relatively: RAF Family, 0.282 to 6.2 and 0.185 to 3.030, Clark Y Family, 0.318 to 4.530 and 0.167 to 3.290, Reed airfoil 0.698 and 0.435, Flat Plate 3.08 and 0.42. (L. J. Briggs and H. L. Drygen, "Aerodynamic characteristics of twenty-four airfoils at high speeds," N. A. C. A. Tech. Report No. 319, Washington, D. C.) The maximum theoretical velocity of the exhaust gases is $c_t=9610(Q)^{\frac{1}{2}}$ ft./sec., where Q denotes the number of kilocalories per gram of mixture. The value of Q for several substances is: acetaldehyde 1.32, benzene 2.46, methane 2.63, normal hydrogen and normal oxygen 3.21, normal hydrogen and monoatomic oxygen 7.73, monoatomic hydrogen and monoatomic oxygen 13.35. (Aldo Vieira Da Rosa, "Notes of rocket fuels," J. Am. Rocket Soc., No. 61, 4–6 (March, 1945).) The hydrogen oxygen mixtures with Q greater than 3.21 are very unstable and they have not, as yet, been used.

The values of α and c_t may refer to gasoline. The result is

W_2/W_1	1:50	1:100	1:200	1:300
(s_2-s_1) in miles	1897	2220	2558	2745

As is seen, long distance bombing or mailing demands a high mass ratio. For $c=30{,}000$ ft./sec., $\alpha=15$, $\epsilon=\frac{1}{5}$, one obtains

W_2/W_1	1:50	1:100	1:200	1:300
(s_2-s_1) in miles	25,900	30,380	34,800	37,420

A final remark should be made concerning formulas (7) and (27) given here in the original form in which they appeared in the literature. These formulas were derived for h in meters and v in m/sec. For calculations in ft. and ft./sec. one has to change the denominator of the second term inside the bracket in (7), i.e., to multiply it by 3.28083 and the numerator of the first term in (27), i.e., to multiply it by 10.763845. All the calculations based on these formulas should be adjusted to the new dimensions, i.e., the numerical coefficients in those formulas will be subject to a change.

Design of Broad Band I.F. Amplifiers

RICHARD F. BAUM
Raytheon Manufacturing Company, Waltham, Massachusetts
(Received February 19, 1946)

This paper deals with the design of broad band I.F. amplifiers with single tuned circuits in t stages. It is found that the figures of merit (Q_n) of the individual circuits should be related to the Q of a reference circuit according to:

$$Q/Q_n=\sin\left[(2m+1)\pi/2t\right]\cdot\frac{\omega_0}{\omega_{0n}},\quad m=0,\ 1,\ 2,\ \cdots(t-1).$$

Then by proper tuning either an oscillatory or a monotonic response may be obtained. The relative band width BW/f_0 and the gain tolerance d_0 within the band determines the value of Q. The minimum number of stages for a given minimum attentuation in the cut-off region depends only on the gain tolerance and on the desired kind of response. Gain maxima (attenuation minima) appear at frequency deviations $\Delta f_n{}^{\min}$ from middle band frequency f_0 given by:

$$2\Delta f_n{}^{\min}/BW=\cos\left[(2m+1)\pi/2t\right].$$

Their location depends only on the number of stages. The resonance deviations Δf_{0n} of the tuned circuits are proportional to $\Delta f_n{}^{\min}$ with a proportionality factor F dependent on t and d_0. The circuit impedances are calculated from a prescribed gain or from the maximum attainable gain. A formula for the maximum gain band width product is derived. The result is extended to double tuned stages. A practical example is given.

(A) INTRODUCTION

OF several types of broad band I.F. amplifiers the types designated as "stagger tuned" will be considered here. They comprise a number of single tuned stages with distinct resonance frequencies and thus differ from the normal I.F. amplifier group using identical stages. The design problem is mainly the proper distribution of the resonance frequencies within the frequency band and the attribution of proper selectivities to the individual circuits.

Amplifiers of this type have been described by several authors and their results given for a restricted number of circuits in tabulated form. It is obvious that a design procedure derived from general formulas is preferable, provided they do not include too complicated functions which obscure the mutual dependence of the parameters. One advantage of the proposed design consists in the simplicity of the resulting formulas, which involve only trigonometric and hyperbolic functions of real variables.

The success of the method is based on the very common assumption that the detuning ϵ of a resonant circuit is proportional to the frequency deviation from the resonance frequency f_0:

$$\epsilon = 2(f-f_0)/f_0,$$

whereas the exact expression reads:

$$\epsilon = f/f_0 - f_0/f.$$

This approximation restricts the correctness of all derived formulas to a limited range, because for a band width of about one-half the center frequency the error in the gain of the stages resonating at one end of the band reaches 10 percent at the other end. A slight detuning of the outside stages is needed for compensation. With a band width of one-quarter the mid-band frequency this error could be compensated. In the case of wider bands the method still furnishes a first approach to the correct design values.

With this approximation the response curve of the amplifier becomes a function expressed by a polynomial of even powers of ϵ. On the other hand, the most desirable response curve may be developed in a polynomial of the same power of ϵ. By equalizing the two polynomials all circuit parameters may be calculated. This method of attack was outlined by Feldtkeller[1] who used the method of equalizing the coefficients of equal powers of ϵ. His method leads to a system of non-linear equations, the solution of which seems impossible for more than four circuits. Furthermore, it does not lead to general design formulas.

If the validity of the polynomials is continued into the region of complex frequencies then they can be equalized by equalizing their complex roots. This method has been applied in the present case. Use is made of equivalent representation of the response curve derived from the Tschebyscheff function. The roots of this expression can be calculated easily and lead to a very comprehensive result.

Considered as given design quantities are the ratio of band width to mid-band frequency, the gain, the gain tolerance within the band, and a certain rate of cut-off. Alternatively the design can be based on maximum gain.

[1] Dr. R. Feldtkeller, *Theorie der Rundfunksiebschaltungen* (S. Hirzel, Leipzig).

(B) POSSIBLE ATTENUATION CURVES

The following symbols are used:

g_n = mutual conductance,
R_n = shunt resistance,
Q_n = figure of merit,
ϵ_n = twice the relative detuning,
$\omega = 2\pi$ times the frequency f,
$\omega_0 = 2\pi$ times the mid band frequency f_0,
$\omega_{0n} = 2\pi$ times the resonant frequency f_{0n},
s_n = normalized detuning,
Δs_{0n} = normalized resonance displacement,
s_L = normalized half-band width,
t = number of tuned circuits,
BW = band width, Δf_0 = half the band width } according to definition,
d = attenuation (Nepers),
d_0 = gain tolerance within the band (Nepers),
d_m = minimum required attenuation outside the band (Nepers),
l = loss in amplification,
$m = 0, 1, 2, \cdots (t-1)$,
D_P = possible response curve,
D_r = required response curve, and
G = gain.

The index n refers to the nth circuit of the amplifier. Without this index the quantities apply to a reference circuit.

The gain of a single tuned stage is

$$G_n = g_n R_n/(1+jQ_n\epsilon_n), \tag{1}$$

$$|G_n| = g_n R_n \cdot (1+Q_n^2\epsilon_n^2)^{-\frac{1}{2}}, \tag{2}$$

$$= g_n R_n \cdot (1+s_n^2)^{-\frac{1}{2}}. \tag{3}$$

Each tuned circuit is adequately defined by its shunt resistance, its Q_n and resonant frequency f_{0n}. If a number of circuits are considered, we may calculate these quantities for a reference circuit, which does not form part of the amplifier itself, and the ratio of the Q's and the resonance frequency displacement between the circuit in question and the reference circuit. Then the expression for the gain becomes

$$|G_n| = g_n R_n[1+(a_n's+b_n)^2]^{-\frac{1}{2}}. \tag{4}$$

The additional coefficient a_n' defines the Q ratio and the coefficient b_n relates the resonant frequencies. The following relations exist:

[2] Compare: W. W. Hansen, J. App. Phys. **16**, 529 (Sept. 1945); H. Wallman, "Stagger tuned I.F. amplifiers," M.I.T. Report 524 (Feb. 1944).

$$Q_n = R_n/\omega_{0n}L_n, \qquad Q = R/\omega_0 L; \tag{5}$$

$$\epsilon_n = 2(\omega-\omega_{0n})/\omega_{0n}, \qquad \epsilon = 2(\omega-\omega_0)/\omega_0; \tag{6}$$

$$s_n = Q_n\epsilon_n, \qquad s = Q\epsilon; \tag{7}$$

$$a_n = Q_n/Q, \qquad a_n' = a_n\omega_0/\omega_{0n}; \tag{8}$$

$$b_n = a_n\Delta s_n; \tag{9}$$

$$\Delta s_n = Q2(\omega_0-\omega_{0n})/\omega_{0n}. \tag{10}$$

All the detunings and frequency displacements appear in normalized form that is multiplied by Q. The reference circuit is assumed to resonate at the mid-band frequency f_0.

In accordance with Fig. 1, the index $n=1$ and t is reserved to the stages resonating at the edges of the band, the index 2 and $(t-1)$ to the stages next to it, and so on. In the case of an uneven number of stages, the index $(t+1)/2$ will apply to the stage resonating at mid-band frequency.

The gain of t stages is

$$|G| = \prod_{n=1}^{n=t} g_nR_n[1+(a_n's+b_n)^2]^{-\frac{1}{2}}. \tag{11}$$

In the middle of the band s equals zero and the gain becomes

$$|G_0| = \prod_{n=1}^{n=t} g_nR_n[1+b_n^2]^{-\frac{1}{2}}. \tag{12}$$

If all stages would tune at the same frequency, the gain would be

$$|G_0^0| = \prod_{n=1}^{n=t} g_nR_n.$$

The root factor thus defines a loss l_0 in gain because of the displacement of resonant frequencies

$$l^2 = \prod_{n=1}^{n=t} (1+b_n^2) \tag{13}$$

which makes

$$|G_0| = (1/l)\prod_{n=1}^{n=t} g_nR_n, \tag{14}$$

$$|G| = l|G_0|\prod_{n=1}^{n=t}[1+(a_n's+b_n)^2]^{-\frac{1}{2}}. \tag{15}$$

The factor in parenthesis defines the attenuation

$$d = .5\lg\prod_{n=1}^{n=t}[1+(a_n's+b_n)^2], \tag{16}$$

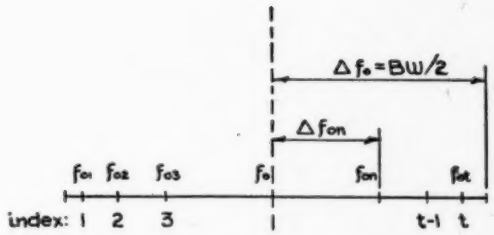


FIG. 1. Disposition of indexes and resonant frequencies. The lowest index $n=1$ is given the stage resonating at the lowest frequency f_{01}. The indexes increase with increasing resonance frequencies.

from which all possible response curves D_P follow with

$$D_P = \epsilon^{2d} = \prod_{n=1}^{n=t}[1+(a_n's+b_n)^2]. \tag{17}$$

D_P refers to the power response, whereas d refers to voltage attenuation and is measured in Nepers.

For reasons of symmetry with respect to the mid-frequency f_0 we anticipate the circuits to be symmetrically disposed with respect to f_0, or

$$a_n' = a_{t-n+1}', \quad b_n = -b_{t-n+1}. \tag{18}$$

In this case the attenuation curve D_P is a polynomial in s containing only even powers of s and of a highest power $k_m = 2t$. This polynomial thus represents the general expression for all possible symmetrical response curves.

(C) REQUIRED RESPONSE CURVE

a. The Oscillatory Form

The required response curve has either an oscillatory form as in Fig. 2a and 2b, or a monotonic form as in Fig. 2c.

The former is defined by an oscillation of the gain between two fixed values. The magnitude of this oscillation shall be called "gain tolerance" d_0 and expressed in Nepers. The latter form shows a steady increase in attenuation without any oscillation. The required response curves are designated D_r.

Each minimum of the oscillatory response curve corresponds to one tuned circuit. Thus it shows t minima and $t-1$ maxima and can be represented by a polynomial of a highest power $k_m = 2t$ of the independent variable. Only even powers appear for reason of symmetry.

The band width shall be defined in a somewhat unusual way by that frequency deviation at

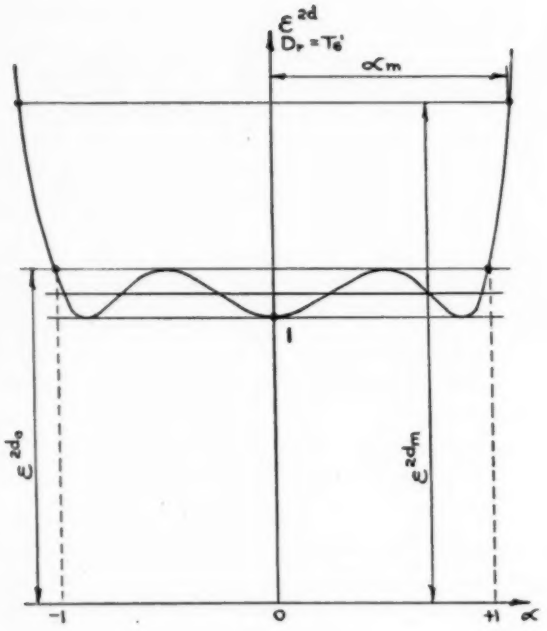


FIG. 2a

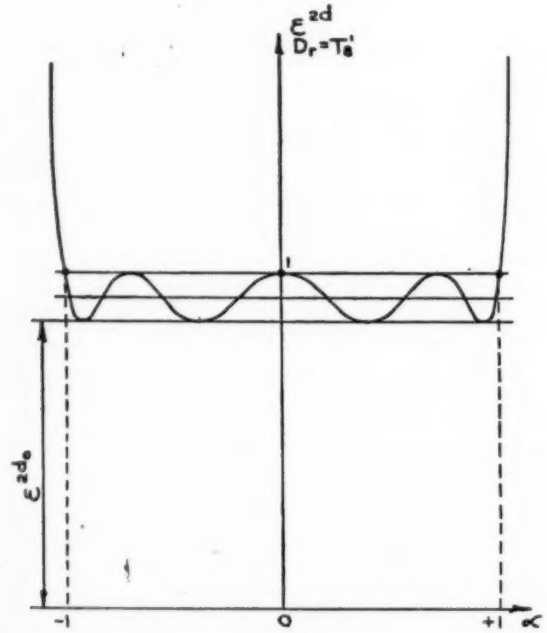


FIG. 2b

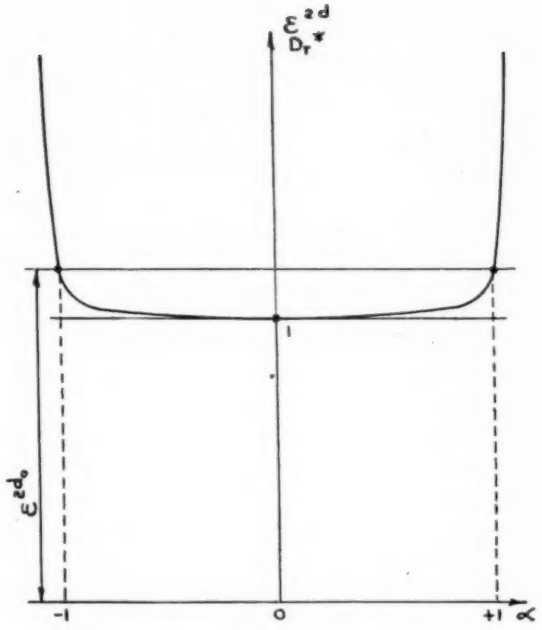


FIG. 2c

FIG. 2. Required response curves *a* and *b* show oscillatory, *c* monotonic behavior. The band width is reached for a prescribed attenuation d_0. The tolerance d_0 is positive for an uneven number of stages (three in Fig. 2a) and negative for an even number (four in Fig. 2b).

which the gain falls off a given amount of d_0 Nepers. In the oscillatory case d_0 coincides with the admissible gain tolerance. This band width will usually be narrower than the one defined by the half-power points.

With a normalized detuning s as abscissa

$$s = Q2\Delta f/f_0$$

the half-band width appears in normalized form as

$$s_L = QBW/f_0. \tag{19}$$

In Fig. 2 the ratio

$$\alpha = s/s_L = 2\Delta f/BW \tag{20}$$

is taken as the independent variable; the band width then is reached for $\alpha = \pm 1$. At mid-frequency $\alpha = 0$ and the response equals one.

It is interesting to note that the polynomial of α of the $(2t)$th order is completely defined by the demanded oscillatory behavior within the band $-1 < \alpha < +1$. Thus there is one definite position of the extremes within the band.

As proof, the roots of the derivative equation $dD_r/d\alpha$ may be extracted. They are functions of $(2t)$ coefficients. If introduced into the equation for D_r and if D_r is equalized to the prescribed extreme values, a system of $(2t-1)$ equations is found for $(2t+1)$ unknown coefficients. The remaining two equations derive from a prescribed value D_{r0} at the ends of the band.

If we can find another polynomial T_k' of α which within the same band shows the same oscillatory behavior as the polynomial D_r, the two have to be identical. The reason for introducing a new function T_k' will be apparent, when in the course of calculation the roots of $D_r=0$ have to be known. The roots of an equation of higher degree than the fourth may be expressed in general form only in special cases. With the aid of the new function T_k' it can be shown that the polynomial D_r belongs to such a special case. Furthermore T_k' allows writing down the coefficients of D_r almost without any calculation. Now these coefficients shall be given only as matter of interest; they are not useful for the present purpose.

The function T_k' is derived from the Tschebyscheff function T_k.[3]

$$T_k=\cos\left[k\cos^{-1}\alpha\right], \tag{21}$$

in which k is a parameter. Figure 3 shows this function for $k=6$. T_k oscillates between $+1$ and -1 within a band of $-1<\alpha<+1$. It equals one for $\alpha=\pm1$. The number of minima equals $k/2$. For values of $|\alpha|>1$ it grows steadily and an equivalent expression in this region is

$$T_k=\cosh\left[k\cosh^{-1}\alpha\right]. \tag{22}$$

Within the band, T_k can be best described as a projection onto the $\alpha-T_k$ plane of a cos curve inscribed on the surface of a circular cylinder of unit radius with the T_k axis as its axis, k being the number of periods around the circumference. This explains the location of maxima, minima, and zeros according to Fig. 3 where the circumference of the unit circle is divided into $2k=12$ parts and the points thus found projected onto the α axis. The minima of T_k thus occur at

$$\alpha^{\min}=\cos\left[(2m+1)\pi/k\right]. \tag{23}$$

In order to show that T_k is nothing but a polynomial of α of a highest power α^k, it is best to use a recursion formula obtained as follows:

$$T_k=\cos(ky),\quad y=\cos^{-1}\alpha,$$

$$\cos\left[(k+1)y\right]+\cos\left[(k-1)y\right]=2\cos(ky)\cos y,$$

$$T_{k+1}+T_{k-1}=2T_kT_1. \tag{24}$$

This formula allows one to compute T_{k+1} if T_k and T_{k-1} are known. Now as $T_0=1$ and $T_1=\alpha$ we obtain immediately

$$T_2=2\alpha^2-1,\quad T_3=4\alpha^3-\alpha,$$
$$T_4=8\alpha^4-8\alpha^2+1,\quad \text{etc.}$$

The coefficients obey the law

$$A_{k,m}=2A_{k-1,m-1}-A_{k-2,m}, \tag{25}$$

where $A_{k,m}$ is the coefficient of α^m in the polynomial T_k. This follows from (24).

The desired response curve D_r, Fig. 2a, differs from T_k only by a reduced magnitude of oscillation and the fact that D_r always equals $+1$ for $\alpha=0$, provided we make k equal to twice the number of stages. Thus an expression of the form

$$T_k'=pT_k+q,\quad 0<p<1, \tag{26}$$

$$k=2t \tag{27}$$

can be made identical with D_r by proper choice of the two coefficients p and q. Neither the position of the extremes nor the power of the polynomial is affected.

The magnitude of oscillation is restricted by the maximum allowed gain variation within the

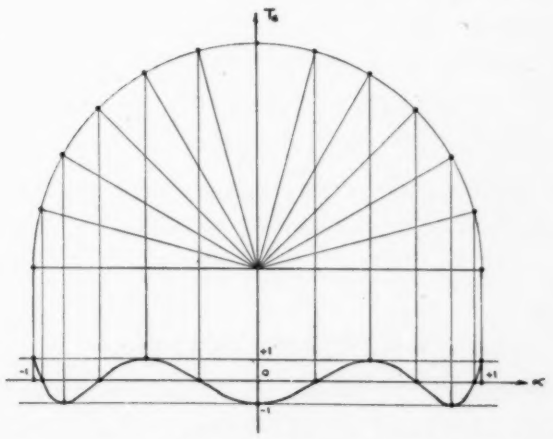


FIG. 3. The Tschebyscheff function $T_k(\alpha)$ for $k=6$. The position of maxima, minima, and zeros is evident from this construction. The number of minimas equals $k/2$. $T_k(\alpha)$ oscillates between $+1$ and -1 and reaches $+1$ for $\alpha=\pm1$.

[3] See for instance, E. Madelung, *Die mathematischen Hilfsmittel des Physikers* (Dover Publications, New York, 1943).

band. It shall be called gain tolerance d_0 (Nepers). Now there is a slight difference in T_k according to the value of $k=2t$, as for an even number of stages

$$T_k(0)=+1$$

and for an uneven number of stages

$$T_k(0)=-1.$$

The requirement then is

$$T_k'(0)=1 \quad \text{for} \quad T_k(\alpha)=T_k(0)$$

and

$$T_k'=\exp(2d_0)\begin{cases}\text{for } T_k=+1 \text{ if it is even,}\\ \text{or } T_k=-1 \text{ if it is uneven.}\end{cases}$$

Furthermore, for an even number of stages the attenuation falls below zero level and d_0 has to be taken negative. The two cases are compared in Fig. 2a and b for $2t=6$ and $2t=8$, respectively. As we consider only small values of d_0 we may set

$$\exp(2d_0)=1+2d_0 \tag{28}$$

obtaining

$$q=1+d_0, \quad p=|d_0|, \tag{29}$$

$$D_r=T_{2t}'=|d_0|\,T_{2t}(\alpha)+(1+d_0). \tag{30}$$

As mentioned before, the purpose of writing the polynomial f or the required attenuation by means of the Tschebyscheff function was the possibility of expressing the roots of $D_r=0$ by a general formula. As these roots are complex, it has to be assumed that (27) is valid within the complex region if it is defined on the real α-axis. The proof of this shall not be given here.

$$D_r=|d_0|\cos[2t\cos^{-1}\alpha_0]+(1+d_0)=0,$$
$$2t\cos^{-1}(\alpha_0)=\cos^{-1}[-(1+d_0)/d_0]. \tag{31}$$

The value of $(1+d_0)/d_0$ is always larger than one. Let an auxiliary constant c be defined by

$$\cosh c=(1+d_0)/d_0 \tag{32}$$

or

$$\cos(jc-\pi)=-(1+d_0)/d_0.$$

Because of the periodicity of the cos function the real part of its argument is multiple valued with a period of 2π:

$$\cos(jc-\pi-2m\pi)=-(1+d_0)/d_0, \qquad m=0,\,1,\,2,\,\cdots.$$

This introduced into Eq. (31) gives for the roots of D_r:

$$\alpha_0=\cos\left[j\frac{c}{2t}-\frac{2m+1}{2t}\pi,\right] \tag{33}$$

and for their real and imaginary part

$$\alpha_R=\cosh(c/2t)\cos[(2m+1)\pi/2t], \tag{34}$$

$$\alpha_I=\sinh(c/2t)\sin[(2m+1)\pi/2t]. \tag{35}$$

b. The Monotonic Form

If in the polynomial defined by Eq. (17) all coefficients but one are made equal to zero, the polynomial reduces to an expression of the form

$$D_P^*=A+Bs^k.^4$$

This suggests the second form of response curve D_P as sketched in Fig. 2c, which shows a flat base within the band. In order to obtain the flattest response together with the steepest cut-off slope, k has to be chosen as large as possible, that is equal to $2t$.

The half-band width in this case shall be defined by the value of $s=s_L$ at which the attenuation attains a prescribed value d_0 (Nepers). Again $\alpha=s/s_L$ is used as the variable and zero attenuation is required at mid-band frequency ($\alpha=0$), which gives with the approximation of Eq. (28)

$$D_r^*=1+2d_0\alpha^{2t}, \tag{36}$$

as the expression for the required monotonic response curve.

The roots of this expression can easily be found

$$\alpha_0^{*2t}=-1/2d_0=\exp[j(\pi+2m\pi)]/2d_0,$$
$$\alpha_0^*=[2d_0]^{-1/2t}\exp[j(2m+1)\pi/2t]; \tag{37}$$

their real and imaginary parts are

$$\alpha_R^*=[2d_0]^{-1/2t}\cos[(2m+1)\pi/2t], \tag{38}$$

$$\alpha_I^*=[2d_0]^{-1/2t}\sin[(2m+1)\pi/2t]. \tag{39}$$

The close relationship of these expressions to Eqs. (34) and (35) should be noted. They contain the same cos and sin factors. The cos factor again coincides with the expression for the location of

[4] All quantities referring exclusively to monotonic response are distinguished by a star. The solution of this case is given in Wallman's paper, reference 2.

the attenuation minima, according to Eq. (23).

$$\alpha^{\min}=\cos\left[(2m+1)\pi/2t\right]. \tag{40}$$

The physical meaning of these resemblances shall be discussed in the next paragraph.

(D) CALCULATION OF CIRCUIT PARAMETERS

a. Selectivities and Resonant Frequencies

By equalizing the roots given by Eq. (37) or Eq. (33) with the roots of Eq. (17), all necessary relationships between circuit parameters are obtained. The roots of Eq. (17) are

$$s_{0n}=b_n/a_n'+j/a_n'; \tag{41}$$

by division with s_L:

$$\alpha_{0n}=b_n/a_n's_L+j/a_n's_L. \tag{42}$$

By comparing Eq. (42) and Eq. (33) we obtain for the oscillatory case

$$1/a_n'=s_L\sinh(c/2t)\sin\left[(2m+1)\pi/2t\right], \tag{43}$$

$$b_n/a_n'=s_L\cosh(c/2t)\cos\left[(2m+1)\pi/2t\right]. \tag{44}$$

The value of s_L now can be fixed by demanding that for an uneven number of circuits the reference circuit should be identical with the circuit resonating at mid-band frequency, that is by requiring $a_n'=1$, $b_n=0$ for $m=(t-1)/2$. This is possible only if:

$$s_L=1/\sinh(c/2t). \tag{45}$$

Together with Eqs. (8) and (9)

$$1/a_n'=(Q/Q_n)(\omega_{0n}/\omega_0)$$

$$=\sin\left[(2m+1)\pi/2t\right], \tag{46}$$

$$b_n/a_n's_L=(\Delta f_{0n}/\Delta f_0)=\alpha^{\min}\cosh(c/2t), \tag{47}$$

where Δf_{0n} is the displacement of the resonance frequency from the mid-band frequency and Δf_0 is half the band width, as shown in Fig. 1.

Between m and n a relation exists of the form

$$m=n-1, \tag{48}$$

which means that by setting successively m equal to $m=0,\ 1,\ 2,\ 3,\ \cdots(t-1)$ all circuit parameters for $n=1,\ 2,\ 3,\ \cdots t$ are obtained.

The same reasoning in the case of a monotonic response curve leads to the formula

$$s_L^*=\left[2d_0\right]^{1/2t}, \tag{49}$$

$$(Q^*/Q_n^*)(\omega_{0n}^*/\omega_0)=\sin\left[(2m+1)\pi/2t\right], \tag{50}$$

$$\Delta f_{0n}^*/\Delta f_0=\alpha^{\min}\cdot\left[2d_0\right]^{-1/2t}. \tag{51}$$

Equation (50) is identical with Eq. (46). For the following discussion, Eq. (40) is joined in the form

$$\alpha^{\min}=\frac{\Delta f_n{}^{\min}}{\Delta f_0}=\cos\left[(2m+1)\pi/2t\right]. \tag{52}$$

It should be noted that, in the first approximation, the factor ω_{0n}/ω_0 and ω_{0n}^*/ω_0 equals one; but that in the case of wide bands and especially for the circuits tuning at the cut-off region, this factor may be considerably different from 1. The asymmetry it introduces shows an increase in Q when passing through the band from the low to the high frequency end.

The best insight into the result is obtained by using the first approximation only and may be stated now this way:

> In order to achieve either a monotonic *or* an oscillatory type of response for a given number of t circuits, the selectivities of the individual circuits have to be staggered according to the law
>
> $$Q/Q_n\leqq\sin\left[(2m+1)\pi/2t\right]. \tag{53}$$

Thus the circuit tuning nearest to the mid-band frequency f_0 receives the lowest Q and the further a circuit tunes off f_0 the higher a Q it obtains.

The circuits of a monotonic amplifier should resonate according to Eq. (51) at frequency displacements proportional to the displacements, where an amplifier with oscillatory response has its gain peaks, that is at resonance displacements

$$\Delta f_{0n}/\Delta f_0=\left[2d_0\right]^{-1/2t}\cos\left[(2m+1)\pi/2t\right]. \tag{54}$$

From this expression it is apparent that the resonance points crowd more and more together as one moves from the middle to the edges of the band. This behavior as well as the increased selectivity could have been predicted from the general appearance of the Tschebyscheff function.

For an amplifier of oscillatory response the resonance displacement is larger by a proportionality factor F

$$F=\cosh(c/2t), \tag{55}$$

where

$$\cosh c=(1+d_0)/|d_0|.$$

The larger the admissible gain variation d_0 and the larger the number of stages, the better the resonance frequencies and the peak frequencies will coincide.

The absolute value of selectivity follows from Eq. (19) with

$$Q = s_L f_0 / BW.$$

For a given relative band width BW/f_0 thus

$$QBW/f_0 = s_L. \tag{56}$$

The larger the relative band width the lower the Q of the circuits has to be. For a given relative band width for oscillatory (Eq. (45)) or monotonic response (Eq. (49)) Q increases with the gain tolerance d_0.

b. The Minimum Number of Circuits

The minimum number of circuits can be calculated if a minimum attenuation of d_m (Nepers) at a given frequency displacement from mid-band frequency is required, as shown in Fig. 2a. If this displacement equals α_m times the half-band width, then with Eq. (22) and (30)

$$|d_0| \cosh [2t \cosh^{-1} \alpha_m] + (1+d_0) \geqq \epsilon^{2d_m},$$

$$\cosh [2t \cosh^{-1} \alpha_m] \geqq [\epsilon^{2d_m} - (1+d_0)]/|d_0|.$$

Introducing an auxiliary quantity h by

$$\cosh h = [\epsilon^{2d_m} - (1+d_0)]/|d_0|, \tag{57}$$

and solving for t

$$2t \geqq h/\cosh^{-1} \alpha_m. \tag{58}$$

The minimum number of stages thus increases with the required attenuation d_m but decreases with the admissible gain variation d_0. In most cases $\epsilon^{2d_m} \gg (1+d_0)$ and Eq. (57) can be reduced to

$$\cosh h \cong \epsilon^{2d_m}/|d_0|, \tag{59}$$

and further to

$$h \cong 2d_m - lg(d_0/2). \tag{60}$$

On the other hand a given attenuation will be reached for frequencies given by

$$\alpha_m' = \cosh (h/2t). \tag{61}$$

The corresponding formulas for the monotonic case, using Eq. (36), are $1+2d_0\alpha_m^{2t} \geqq \epsilon^{2d_m}$

$$h^* = lg[(\epsilon^{2d_m} - 1)/2d_0] \tag{62}$$

and

$$2t^* \geqq h^*/lg\alpha_m. \tag{63}$$

Equation (62) again most frequently reduces to

$$h^* \cong 2d_m - lg(2d_0). \tag{64}$$

Equations (57) to (64) should be used with caution as they refer to the behavior of the response curves outside the band width. It was pointed out before that in this region the assumption for ϵ (according to Eq. (6)) becomes increasingly incorrect. It is well to restrict α_m to values $1 < \alpha_m < 2$. Also if the required minimum attenuation d_m is taken very large compared with the gain tolerance, that is if

$$2d_m \gg lgd_0,$$

the minimum number of stages becomes proportional to d_m. Then it would not make much difference if the amplifier has t stages or if it is broken up, for instance into two identical parts of $t/2$ stages each. This is another reason to restrict α_m to small values, where the attenuation is not so large.

By means of (58) and (63), it can be shown that within the restricted region of α_m breaking up of the amplifier into two or more parts leads to a greater number of stages for the same required attenuation d_m and the same tolerance d_0. This statement, in other words, says that a single unit amplifier designed according to the present scheme has a better cut-off characteristic than any other amplifier of the same number of stages either identical or identical in pairs or triplets or more sub-units. Indeed for the single unit amplifier (with (58) and (60))

$$2t = [2d_m - lg(d_0/2)]/\cosh^{-1} \alpha_m.$$

If this amplifier is composed of i identical parts, then each has to have a minimum attenuation d_m/i and a tolerance d_0/i. The total number of stages necessary is:

$$i \cdot 2t_i = i[2d_m/i - lg(d_0/2i)]/\cosh^{-1} \alpha_m.$$

The previous statement is correct if always

$$i[2d_m/i - lg(d_0/2i)] > 2d_m - lg(d_0/2)$$

or if

$$[d_0/2]^{i-1} < i^i;$$

this is correct, as always

$$i>1 \quad \text{and} \quad d_0<1.$$

The analogous proof in the case of monotonic response leads to the condition:

$$[2d_0]^{i-1}<i^i$$

which is also true.

This result represents a justification for the choice of the required response curves. Another one is derived from the property of the Tschebyscheff function $T_k(\alpha)$ of being the one polynomial of highest exponent k and smallest coefficient $+1$ or -1 which within the band $-1<\alpha<+1$ approaches closest zero. Any other polynomial of the same order k attains a larger value at least at one point within the band. The chosen response curve D_r thus is the curve which within the band most closely approaches constant gain, if the number of circuits is prescribed.

c. Impedances and Gain

The staggering of resonance frequencies results in a loss l in gain at the mid-band frequency given by Eq. (13). By dividing Eq. (44) by Eq. (43), l becomes

$$l^2=\prod_{m=0}^{m=t-1}[1+\coth^2(c/2t) \times\cot^2[(2m+1)\pi/2t]] \quad (65)$$

in the oscillatory case, whereas in the monotonic case

$$l^{*2}=\prod_{m=0}^{m=t-1}[1+\cot^2[(2m+1)\pi/2t]]$$

or

$$l^*=\prod_{m=0}^{m=t-1}[1/\sin[(2m+1)\pi/2t]]=2^{t-1}. \quad (66)$$

Each additional stage here reduces the gain by a factor of 2.

In order to calculate the gain, an assumption about the shunt resistance R_n has to be made. The stages of the amplifier will most likely be designed either with the same capacity C and inductively tuned, or with the same inductance L and capacitively tuned. In the former case

$$Q_n=C\omega_{0n}R_n, \quad Q=C\omega_0R, \qquad R_n/R=Q_n\omega_0/Q\omega_{0n}=a_n'. \quad (67)$$

In the latter case

$$Q_n=R_n/\omega_{0n}L, \quad Q=R/\omega_0L, \qquad R_n/R=Q_n\omega_{0n}/Q\omega_0=a_n'(\omega_{0n}/\omega_0)^2. \quad (68)$$

Only the first case shall be considered here, as it gives a slightly larger gain. The corresponding equations for the other case can be derived easily. Introducing Eq. (67) into the gain formula (14)

$$|G_0|=(g^tR^t/l)\prod_{n=1}^{n=t}a_n' \quad (69)$$

and with

$$\prod_{n=1}^{n=t}a_n'=\prod_{m=0}^{m=t-1}[1/\sin[(2m+1)\pi/2t]], \qquad =L^*=2^{t-1}, \quad (70)$$

$$|G_0|=(2gR)^t/2l. \quad (71)$$

Identical tubes have been assumed. The first stage often consists of an input transformer with a step-up ratio g_T. Then g^t would have to be replaced by g_Tg^{t-1}.

In some cases by reason of the large number of tubes involved the required gain is easily achieved and R may be calculated from Eq. (71)

$$\log_{10}(2gR)=[\log_{10}|G_0|+\log_{10}2l]/t. \quad (72)$$

For the monotic response the gain expression becomes extremely simple because of the relation (70); namely,

$$|G_0^*|=(gR)^t. \quad (73)$$

The gain of the monotonic amplifier is the same as the gain of an amplifier consisting only of reference circuits, all tuned at mid-frequency.

Gain formula (71) involves the loss l, which has to be calculated by means of Eq. (65). It is a function of $c/2t=\gamma$. In order to facilitate the practical design $\log_{10}(2l)$ is drawn in Fig. 4 as a function of t and γ.

The gain can as well be represented as a function of the capacity C. With

$$Q=RC\omega_0$$

and

$$QBW/f_0=s_L$$

the shunt resistance equals

$$R=s_Lf_0/C\omega_0BW \quad (74)$$

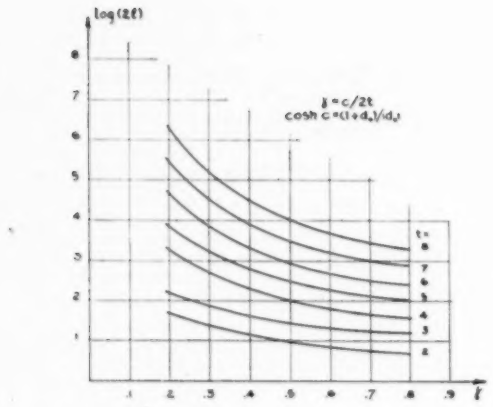


FIG. 4. Value of log $(2l)$ as a function of γ and the number of stages t. The loss in gain l is caused by the staggering of the resonant frequencies.

and the gain

$$|G_0| = [(g/\pi C)(s_L/BW)]^t/2l \tag{75}$$

and

$$|G_0^*| = [(g/\pi C)(s_L^*/2BW)]^t, \text{ respectively.} \tag{76}$$

The maximum gain is obtained by setting C equal to the sum of input and output capacities of the tube

$$C = C_{in} + C_{out} = C_m.$$

The maximum average gain band width product follows with

$$|\bar{G}_0| BW = (g/\pi C_m) s_L (2l)^{-1/t} \tag{77}$$

and

$$|\bar{G}_0^*| BW = (g/\pi C_m) s_L^*/2. \tag{78}$$

In the quantity in parenthesis the theoretical maximum is recognized, obtainable by optimum distribution of an infinite number of circuits within the band.[5] The additional factor then will give a clue for the estimation of the gain efficiency of the presented method. It should not be forgotten though that the theoretical maximum is based on constant gain within the band and zero tolerance for the definition of the band width, whereas for a limited number of circuits the band width has to be defined by a given finite tolerance.[5]

This additional factor μ

$$\mu = s_L (2l)^{-1/t}$$

is within the practical region for γ, namely, $.2 < \gamma < .8$, almost independent of the number of stages and given in Table I, for 3 to 8 stages (with an accuracy of ± 5 percent). We thus find $.81 < \mu < .45$.

In the monotonic case

$$\mu^* = .5(2d_0)^{1/2t}$$

which is always smaller than one-half and approaches this value as the number of stages t increases.

(E) EXTENSION OF RESULTS

Instead of using only one single tuned circuit in each stage, we may use two or more tuned circuits either capacitively or magnetically coupled. With certain assumptions all of them lead to response curves represented by polynomials of ϵ. This equivalence is treated in a series of articles.[6] The results obtained for stagger tuned single circuits apply just as well to these cases provided the calculated coefficients are properly interpreted.

Of special interest is the case of two magnetically coupled circuits. If the coupling is specified by a coupling coefficient n (equal to one for transitional coupling) the gain of one stage can be written

$$\begin{aligned} |G_n'| &\leqq [(1+n_n^2-s^2)^2+4s^2]^{-\frac{1}{2}} \\ &\leqq [1+(s+n_n)^2]^{-\frac{1}{2}}[1+(s-n_n)^2]^{-\frac{1}{2}}. \end{aligned} \tag{79}$$

By comparing with the gain of two single tuned stages

$$|G_n''| = g^2 R_n^2 [1+(a_n' s+b_n)^2]^{-\frac{1}{2}}[1+(a_n' s-b_n)^2]^{-\frac{1}{2}},$$

it follows that the coefficient b_n now has to be interpreted as coupling coefficient. The distribution of coupling through the band thus is given by

$$n_n = \coth(c/2t) \cot[(2m+1)\pi/2t] \tag{80}$$

and

$$n_n^* = \cot[(2m+1)\pi/2t]. \tag{81}$$

Furthermore $a_n' = 1$ requires $RC = R_n C_n$. The number of stages now equals $2t$.

[5] See W. W. Hansen, reference 2.

[6] Maynard, Communications 38 (Jan. 1945); Gardiner and Maynard, Proc. I.R.E. 676 (Nov. 1944); R. Feldtkeller, reference 1.

(F) EXAMPLE

An example best shows the application of the derived formulas to a practical problem:

An intermediate frequency stagger tuned amplifier is to be designed having a mid-band frequency of 2 mc, a band width of .5 mc, and a gain of 100,000. The gain tolerance is $d_0=.1$ Nepers and an attenuation of at least 20 db (or 2, 3 Nepers) has to be reached at 1, 3 times the band width. An oscillatory response is desired. Thus:

$$f_0=2 \text{ mc}, \quad \Delta f_0=.25 \text{ mc}, \quad |G_0|=10^5,$$

$$d_0=.1, \quad d_m=2.3, \quad \alpha_m=1.3.$$

$$h=4.60+.693+2.302=7.595, \qquad (60)$$

$$2l \geqq 7.595/.76=10; \qquad (58)$$

thus at least 5 stages are needed. Maximum gain occurs at:

$$\Delta f_n^{\min}/\Delta f_0=\pm.951 \text{ and } \pm.588 \text{ and } 0. \qquad (52)$$

$$\Delta f_n^{\min}=\pm.238 \text{ and } \pm.147 \text{ and } 0 \text{ mc.}$$

The correction factors in (50) and (51) thus are approximately

$$f_{0n}/f_0 \cong f_n^{\min}/f_0=.88, \ .926, \ 1, \ 1.073, \text{ and } 1.119.$$

The auxiliary parameters c and s_L give the selectivities:

$$c=\cosh^{-1}\frac{1.1}{.1}=\cosh 11=3.09, \qquad (32)$$

$$\gamma=c/2l=.309, \quad s_L=3.18, \qquad (45)$$

$$Q=3.18\frac{2}{.5}=12.72, \qquad (56)$$

$$\frac{Q}{Q_n}\frac{\omega_{0n}}{\omega_0} \cong \frac{Q}{Q_n}\frac{f_n^{\min}}{f_0}=.309 \text{ and } .809 \text{ and } 1. \qquad (46)$$

$$Q_1=36.2, \quad Q_2=14, \quad Q_3=12.72=Q,$$

$$Q_5=46.1, \quad Q_4=16.3.$$

The resonance displacements Δf_{0n} derive from

TABLE I. Value of μ for 2 to 8 stages for different values of γ. The maximum gain theoretically attainable corresponds to $\mu=1$.

	$\gamma=c/2l=$					
.2	.3	.4	.5	.6	.7	.8
$\mu=.81$	.75	.675	.61	.55	.5	.45

$\cosh(c/2l)=\cosh .309=1.048,$

$$\frac{\Delta f_{0n}}{\Delta f_0}=\pm.998 \text{ and } \pm.617 \text{ and } 0. \qquad (47)$$

$$\left.\begin{aligned}\Delta f_{01}&=-.250, \quad \Delta f_{02}=-.154, \quad \Delta f_{03}=0\\ \Delta f_{05}&=+.250, \quad \Delta f_{04}=+.154.\end{aligned}\right\} \text{ mc.}$$

The shunt resistances follow from the required gain:

$$\log_{10}(2gR)=[5+3.19]/5=1.64, \qquad (72)$$

where $\log(2l)$ is taken from Fig. 4. With a transconductance $g=5000$ μmhos.,

$$R=4370 \text{ ohms.}$$

$$R/R_n=1/a_n'=.309 \text{ and } .809 \text{ and } 1. \quad (67) \text{ and } (46)$$

$$\left.\begin{aligned}R_1=R_5&=10900\\ R_2=R_4&=5300\\ R_3=R&=4370\end{aligned}\right\} \text{ ohms.}$$

Each circuit thus is specified by its shunt resistance, its Q and its resonance point.

(G) CONCLUSION

A method has been outlined for the design of broad band amplifiers, based on identification of the response curve with either the Tschebycheff polynomial or a parabola of higher degree. The former applies to an oscillatory and the latter to a monotonic response. The simplicity of the resulting formulas was achieved by an approximation which restricts their use to a limited range of band width. The method of attack lends itself just as well to the exact calculation (for any band width) and shall be published in another paper.

The Relation between Nodal Positions and Standing Wave Ratio in a Composite Transmission System*

EUGENE FEENBERG
New York University, New York, New York
(Received December 21, 1945)

Reflection generally occurs at a lossless transition region joining two uniform lossless lines. If the output line feeds into a matched load (no reflection) a standing wave ratio η_0 different from unity exists on the input side of the transition region. If the output line is terminated by a movable short circuit, a relation exists between the nodal positions on opposite sides of the transition section. The relation can be used to determine η_0 thus dispensing with the need for a calibrated detecting system to measure this quantity.

FOR the purpose of this note a composite transmission system is defined to be a transmission line composed of two uniform lossless sections (coaxial or wave guide) joined by any sort of lossless adapter. In particular the system may be simply a single uniform line containing a lossless irregularity (window, supporting bead, stub section, etc.) localized somewhere along the line. Systems of this general type are employed in many varieties of microwave equipment and measuring apparatus. Actual systems are not, of course, lossless, but it is generally true that power losses are kept to a minimum by good design.

In Fig. 1 the incident wave travels from left to right. Region 1 feeds into a matched load so that no reflection occurs from the output end of the system. However, reflection may occur at the adapter or irregularity resulting in nodes and a standing wave ratio η_0 different from unity in region 2.

If region 1 is terminated by a shorting plug the location of the nodes in both regions depends on the position of the shorting plug. We show that the standing wave ratio η_0 which exists in region 2 when region 1 feeds into a matched load (no reflection) can be expressed simply in terms of the relation between the nodal positions in the two regions when region 1 is terminated by a shorting plug. This statement has the practical consequence that standing wave ratios can be evaluated without the use of a calibrated detecting system.

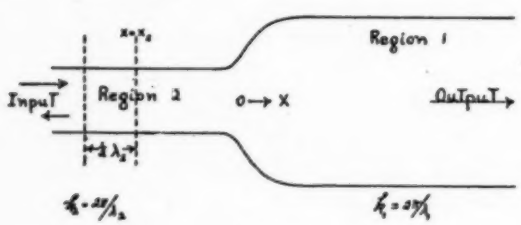


FIG. 1.

The analysis is based on the assumption that outside of a transition region centered about the origin the nodes are uniformly spaced. In each uniform section the distance between alternate nodes defines wave-lengths λ_1 in region 1 and λ_2 in region 2 which may or may not differ from each other and from the wave-length λ in free space. The assumption of uniformly spaced nodes does not exclude the possibility that the uniform sections are capable of transmitting more than one mode without attenuation, but does require that under the experimental conditions only one such mode is present in each uniform section.

Let E represent the voltage (in the case of a coaxial line) or the transverse electric field (in the case of a wave guide). Outside of the transition region the fields are given by the expressions

$$E_1 = A \exp\{-jk_1(x-x_1)\},$$
$$E_2 = B[\exp\{-jk_2(x-x_2)\} - b\exp\{jk_2(x-x_2)\}] \quad (1)$$

for the case pictured in Fig. 1. Here A and B are real quantities which may depend on transverse coordinates, but are independent of x. The reflected amplitude b is real and positive. Then the voltage squared standing wave ratio in region 1 has the value

$$\eta_0 = (1+b)^2/(1-b)^2. \quad (2)$$

* Based on reports written for the Research Laboratory of the Sperry Gyroscope Company, Garden City, New York.

Figure 2 illustrates the field relations when region 1 is terminated by a shorting plug.

To construct solutions for this case we make use of the fact that the complex conjugate of the field defined by Eq. (1) is also a possible field (the associated magnetic field or current quantity is replaced by the negative of its conjugate when E is replaced by E^*). Taking a suitable linear combination of the two fields at our disposal we obtain

$$E_1' = A[\exp\{-jk_1(x-x_1')\} - \exp\{jk_1(x-x_1')\}],$$
$$E_2' = B\exp\{-jk_2(x-x_2)\}[\exp\{jk_1(x_1'-x_1)\} + b\exp\{-jk_1(x_1'-x_1)\}]$$
$$-B\exp\{jk_2(x-x_2)\}[\exp\{-jk_1(x_1'-x_1)\} + b\exp\{jk_1(x_1'-x_1)\}]. \quad (3)$$

This solution vanishes at the short ($x=x_1'$) and thus represents the situation pictured in Fig. 2. In region 2 a nodal surface is located at the position

$$x_2' = x_2 + 1/k_2 \cdot \tan^{-1}[1/\eta_0^{\frac{1}{2}} \cdot \tan k_1(x_1'-x_1)]. \quad (4)$$

Equation (4) can be transformed into

$$k_2(x_2'-x_2) = k_1(x_1'-x_1) - \tan^{-1}\left[\frac{(\eta_0^{\frac{1}{2}}-1)\tan k_1(x_1'-x_1)}{\eta_0^{\frac{1}{2}} + \{\tan k_1(x_1'-x_1)\}^2}\right]. \quad (5)$$

The Eqs. (4) and (5) reduce to the linear relation

$$k_2(x_2'-x_2) = k_1(x_1'-x_1) \pm n\pi, \quad (6)$$

if the transition region produces no reflection ($\eta_0 = 1$).

A plot of the quantity $k_2x_2' - k_1x_1'$ is shown in Fig. 3. The range D over which $k_2x_2' - k_1x_1'$ varies is readily found to have the value

$$D = 2\tan^{-1}\tfrac{1}{2}(\eta_0^{\frac{1}{2}} - \eta_0^{-\frac{1}{2}}) \cong \eta_0^{\frac{1}{2}} - \eta_0^{-\frac{1}{2}} \cong 2b \text{ (when } b \ll 1). \quad (7)$$

The maximum and minimum values of $k_2x_2' - k_1x_1'$ occur at values of x_1' given by the relation

$$\tan k_1(x_1'-x_1) = \pm\eta_0^{\frac{1}{2}}. \quad (8)$$

If η_0 is known, Eq. (7) supplies the answer to the question of how much the relation between the nodal displacements departs from linearity. Conversely an experimental determination of the relation between the nodal displacements yields a value for η_0.

The phase k_1x_1 which appears in Eq. (1) is not readily accessible to measurement under the conditions of Fig. 1. However, a suitable linear combination of Eq. (1) and its conjugate complex yields a solution for the situation in which input and output are interchanged and no reflected wave occurs in region 2. This solution is

$$E_2'' = KB(1-b^2)^{\frac{1}{2}}\exp\{jk_2(x-x_2)\}$$
$$E_1'' = KA(1-b^2)^{-\frac{1}{2}}[\exp\{jk_1(x-x_1)\} + b\exp\{-jk_1(x-x_1)\}], \quad (9)$$

in which K is a constant. Thus nodes occur in region 1 at the positions

$$k_1x_1'' = k_1x_1 + (n_1 - \tfrac{1}{2})\pi, \quad n_1 = 0, \pm1, \cdots \quad (10)$$

when the source is located on the right (region 1) and region 2 feeds into a matched load (no reflection).

In the special case of a uniform line containing a reflecting element (which may occupy an arbitrary length of the line) Poynting's theorem yields the relation

$$A^2 = B^2(1-b^2). \quad (11)$$

With the help of Eq. (11) the double primed solution can be expressed in the form

$$E_2'' = KA\exp\{jk_1(x-x_2)\},$$
$$E_1'' = KB[\exp\{jk_1(x-x_1)\} + b\exp\{-jk_1(x-x_1)\}]. \quad (12)$$

Now, assuming a *symmetrical* reflecting element, Eqs. (1) and (12) are related in the same way as object and image in a plane mirror. Applying the mirror image relation to the location of the

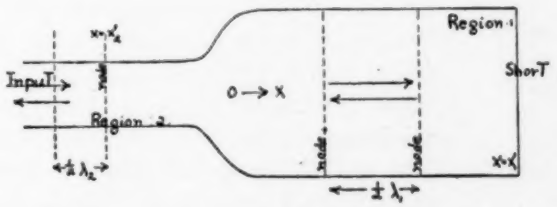


FIG. 2.

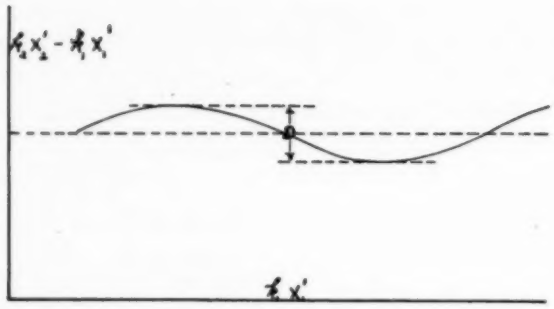


FIG. 3.

nodes we get

$$k_1x_1'' = k_1x_1 + (n_1 - \tfrac{1}{2})\pi = -k_1x_2 + n_2\pi, \quad n_2 = 0, \pm 1, \cdots. \quad (13)$$

Consequently,

$$k_1(x_1 + x_2) = (n + \tfrac{1}{2})\pi, \quad n = 0, \pm 1, \cdots. \quad (14)$$

In the limiting situation produced by reducing the irregularity to the vanishing point continuity requires

$$k_1x_2 = k_1x_1 + m\pi, \quad m = 0, \pm 1, \cdots \quad (15)$$

or

$$k_1x_2 = (2p+1)\pi/4, \quad p = 0, \pm 1, \cdots. \quad (16)$$

Equation (16) can also be expressed in the form

$$x_2 = \lambda_1/8 + p\lambda_1/4. \quad (17)$$

There are evidently two physically distinct cases; one associated with even values of p, the other with odd values. One can show that an infinitesimal lumped capacitative element across the uniform line produces a node at $x_2 = -\lambda_1/8$ (when power flows from left to right). On the other hand a very large lumped inductive element across the line places a node at $x_2 = -3\lambda_1/8$. Thus the lumped capacitative case is associated with odd values of p and the lumped inductive case with even values.

Erratum: Small Prism Infra-Red Spectroscopy

[J. App. Phys. **16**, 77–86 (1945)]

R. BOWLING BARNES, ROBERT S. MCDONALD, AND VAN ZANDT WILLIAMS
Stamford Research Laboratories, American Cyanamid Company, Stamford, Connecticut
AND
RICHARD F. KINNAIRD
The Perkin-Elmer Corporation, Glenbrook, Connecticut

E. GANZ, Basle, Switzerland, has called our attention to an error in our paper. The formula on page 81 should read

$$\Delta\nu(\mathrm{cm}^{-1}) = \frac{\overbrace{\nu^2\left(1 - n^2\sin^2\frac{\alpha}{2}\right)^{\frac{1}{2}}}^{\mathrm{A}}}{8\sin\frac{\alpha}{2}\frac{dn}{d\lambda}}\frac{S_1 + S_2}{f} + \overbrace{\frac{\nu}{2b\dfrac{dn}{d\lambda}}}^{\mathrm{B}}.$$

This was an error in the submitted manuscript. The above formula was used in calculating the results of Table I.

Discussion with Dr. Ganz has brought out the fact that the contribution to the spectral slit width caused by the limiting resolving power of the prism, the B term, is not independent of the slit width, but should be multiplied by a factor which is dependent on the slit width (S). This factor is a maximum of about 0.87 when $S = O$ as shown by P. C. Cross and E. R. Nixon,* and a minimum when $S\ (\mathrm{cm}^{-1}) = \nu/b(dn/d\lambda)$. It remains constant for $S > \nu/b(dn/d\lambda)$. The use of this factor removes the discrepancy between our theoretical and experimental spectral slit widths.

* P. C. Cross and E. R. Nixon, J. Opt. Soc. Am. **34**, 517–520 (1944).

Here and There

New Appointments

Effective March 1, Wallace Waterfall was appointed director of research and product development for the Celotex Corporation. He had been on leave from Celotex for three years, as a member of the scientific staff of Columbia University Division of War Research.

Arthur D. Little, Inc., industrial research organization, has announced two new appointments. Austin W. Fisher will supervise chemical engineering research, and Bruce S. Old will engage in metallurgical research.

Charles D. Coryell, who has been active in atomic research at the Clinton Laboratories, Oak Ridge, Tennessee, has been appointed professor of chemistry at the Massachusetts Institute of Technology. He will join the staff in July.

Carroll L. Wilson, executive assistant since 1942 to Dr. Vannevar Bush, director of the Office of Scientific Research and Development, has been elected vice president of National Research Corporation, Boston, Massachusetts.

Midwest Research Institute, Kansas City, Missouri, recently announced the following new appointments: Max H. Thornton and Loren E. Morey, biochemists, John D. Fulton, biological research specialist, and Howard M. Gadberry, chemist.

J. Mitchel Fain, lately a colonel with the Army Air Forces, has rejoined the staff of Foster D. Snell, Inc., as account executive in charge of technical development for a group of manufacturers of chemical specialties.

John D. Kraus, formerly of the Radio Research Laboratory at Harvard University, has been appointed associate professor in the Department of Electrical Engineering at the Ohio State University.

Effective July 29, Richard A. Beth will become professor of physics and head of the Department of Physics at Western Reserve University. Since 1940 he has been associated with the National Defense Research Committee at Princeton University, directing research on terminal ballistics.

Awards

The annual Young Author's Prize of the Electrochemical Society was presented at the Society's meeting in April to Austin E. Hardy, 25-year-old chemist in the Luminescent Materials Laboratory of the Tube Department of Radio Corporation of America. He received $100 for his paper on "Photoconductivity of Zinc Cadmium Sulphide as Measured with the Cathode-Ray Oscillograph."

The American Russian Institute reports that the 1946 Stalin Prize for Physics was awarded to K. A. Petrzhak and G. N. Flerov for their discovery of spontaneous fission of uranium.

In recent months The Franklin Institute, Philadelphia, Pennsylvania, has presented honor award medals to the following research scientists in recognition of outstanding work in their various fields of interest: Major General Gladeon M. Barnes, Dr. Karl Terzaghi, Dr. George C. Southworth, Dr. Henry C. Sherman, Sir Henry T. Tizard, Dr. Sanford A. Moss, Dr. Bengt Edlén, and Dr. Ira S. Bowen.

New Applied Science Department at Harvard

Harvard University announced in February the formation of a new Department of Engineering Sciences and Applied Physics. The department was organized to consolidate and strengthen the University's research and instruction in applied science and will give instruction at both undergraduate and graduate levels.

A primary educational aim of the new department will be to fill in the "blind spot" which exists between instruction afforded by departments of physics and mathematics in which primary emphasis is upon pure science, and the training afforded by mechanical, electrical, and metallurgical engineering departments in which instruction is concentrated on the technology of engineering practice. The wholesale mobilization during World War II of trained scientists and scientists-in-training taught certain lessons which are being used to guide the planning for the activities of the new department. Administrators and supervisors of many OSRD war research laboratories had an unusual opportunity to observe in action men who had been exposed to various kinds of scientific training. Most of the scientific military objectives centered around novel applications and extensions of principles and techniques which were still in the research laboratories in the late 1930's. It was an almost universal conclusion that those people succeeded best in such advanced development work who had acquired the most thorough basic training in mathematics and fundamental physics, without regard to whether their previous inclinations had veered toward pure science or engineering. The "pure scientists," in particular, were able to adapt their thinking to the requirements of engineering development with surprising facility and success. Since the postwar demands on scientific understanding are likely to become more rather than less exacting, these considerations have influenced the pronounced emphasis on basic training in the programs of study for both undergraduate and graduate degrees.

The undergraduate program will embrace Harvard's series of courses in general education as recommended in the report of the Harvard Committee on Objectives for General Education in a Free Society. It will constitute an "honors" program open only to students qualified to profit by the rigorous training in fundamental mathematics and physics. No attempt will be made to provide at the undergraduate level the various specializations which characterize the different specific branches of engineering. Emphasis will be placed instead on a strong basic training which will qualify the student to pursue graduate work and to carry on research in any of the specialized branches of applied science.

The nucleus of the new Department of Engineering Sciences and Applied Physics is composed of several faculty groups which have been engaged in instruction

and research in acoustics, communications, electronics, applied mathematics, metallurgy, and mechanical engineering science. All of these staff members will bring to their new teaching duties the experience gained from extended war assignments in military research or training where there was opportunity to observe the end results of many methods of instruction in applied science. Associate Professor F. V. Hunt has been appointed chairman of the new department, whose membership will include Professors C. H. Berry and E. L. Chaffee, Associate Professors H. W. Emmons, H. R. Mimno, H. H. Aiken, and R. W. P. King, Assistant Professor L. A. Pipes, and Doctors R. W. Hickman and P. LeCorbeiller. In addition to current endowment support for this group, the University has assigned to the use of the new department the income from $2,000,000 of the Gordon McKay Fund, with which it will be possible to make new appointments.

Among the unique facilities which will be available for research under the auspices of the Department of Engineering Sciences and Applied Physics are the Automatic Sequence Controlled Calculator presented to Harvard in 1944 by T. J. Watson, president of International Business Machines Company, the large anechoic (echo-less) chamber constructed in 1942–43 for free field sound measurements, other rooms specially treated for research in acoustics, and a supersonic wind tunnel constructed at Harvard during the war.

The scholarship program for the department will be patterned after the Harvard National Scholarships. Stipends of these scholarships are adjusted to meet the students' need and provide opportunity for a few carefully selected candidates to pursue their entire program of study without financial concern. Other student aids and teaching fellowships will also be provided.

New A.S.H.V.E. Research Facilities

The American Society of Heating and Ventilating Engineers has purchased property at 7218 Euclid Avenue, Cleveland 3, Ohio, to provide adequate research facilities for serving the entire heating, ventilating, and air conditioning industry as well as the associated industries.

Atomic Research at General Electric

General Electric Company plans an expansion in its program for the study of the basic physics of atomic power, which eventually will require attention of "a substantial part" of its research staff. Already approximately twenty scientists who formerly were engaged on the Manhattan District Project have been appointed to the laboratory since the war's end. The company plans to add another forty technical men to its research staff before the completion of a new laboratory sometime next year.

Expansion at University of Rochester

Recently the Eastman Kodak Company and the Bausch and Lomb Optical Company presented gifts totaling $150,000 to the University of Rochester for its Institute of Optics. This will support the Institute's teaching and research work, and provides for the appointment of Professor Brian O'Brien to the newly created post of research professor of physics and optics.

This enlarged program is in addition to plans for expanding the Department of Physics, for which $550,000 was recently appropriated by the University. The two programs in physics and optics will be coordinated under the administrative direction of Dr. Lee A. DuBridge, chairman of the Department of Physics.

General Electric Science Fellowships

Again this summer General Electric Company will provide all-expense science fellowships at Union College for high school chemistry and physics teachers, selected from ten northeastern States. Fifty teachers will be appointed for the six-week term of study, which will begin July 7.

Electrochemical Society Officers

At the Congress of the Electrochemical Society (first since the beginning of the war) held at Birmingham, Alabama, April 11 to 13, the following new officers were elected:

President	William C. Moore, *New York*
Vice President	A. L. Ferguson, *University of Illinois*
Treasurer	W. W. Winship, *New York*
Secretary	Colin G. Fink, *Columbia University*

New Laboratory for Cluett, Peabody and Company

In April ground was broken for a new $600,000 research laboratory which will carry on scientific research on a scale "without precedent in its industry," according to C. R. Palmer, president of Cluett, Peabody and Company. The company's enlarged research program will be carried out along broad general lines which will include chemical process studies, textile research, mechanical process research, and fundamental research into such projects as the development of new synthetic fibers. The laboratory will contain a number of pilot plants where the results of research developments will be tested in pilot plant operation before being attempted in actual commercial production.

New Quarters for National Research Corporation

Under lease arrangements with Massachusetts Institute of Technology, National Research Corporation will build new quarters on the corner of Memorial Drive and Wadsworth Street, Cambridge, on the Charles River Basin. Construction should be completed by December 15, 1946. Larger facilities are needed as a result of the company's increased research operations and broadening of research programs to include in some fields work for other organizations on a contract basis.

The Bulletin of Mathematical Biophysics

JUNE, 1946

On Oblique Growth of Trees under the Action of Winds: I. OPATOWSKI
Further Contributions to a Probabilistic Interpretation of the Mathematical Biophysics of the Central Nervous System: N. RASHEVSKY
A Note on the Excitation of Nerve Pathways: RICHARD RUNGE
Tree Trunks and Branches as Optimum Mechanical Supports of the Crown: I. The Trunk: MARINUS H. M. ESSER
Outline of a Matrix Algebra for Neural Nets: H. D. LANDAHL AND RICHARD RUNGE
Volume 8, Number 2—The University of Chicago Press, Chicago, Illinois

New Books

General Meteorology

By Horace R. Byers. Pp. x+645. McGraw-Hill Book Company, Inc., New York, 1944. Price $5.

The extensive wartime training of new weather forecasters for the Army and Navy severely taxed the university meteorology teachers. The few available textbooks tended to be either too specialized and advanced for the student's comprehension, or too elementary to teach him enough for his future job. Consequently, many professors wrote self-contained and comprehensive lecture notes at an intermediate level, some of which have recently appeared as books. Because of their origin, these books provide an unusual opportunity for the non-meteorological reader to acquire a very considerable knowledge of weather science without having to refer to numerous specialized texts and articles.

General Meteorology, by Professor Horace R. Byers of the University of Chicago, appears to be the type of book described above, although it is partly based on the same author's 1937 book, *Synoptic and Aeronautical Meteorology*. The great part (in general, the last two-thirds) of *General Meteorology* is devoted to synoptic meteorology—that is, the description of the earth's air masses and their interactions, the use of weather charts in describing and forecasting the weather, the general circulation, and such special topics as tropical cyclones. Also included in this part are five valuable chapters of particular importance to aeronautical meteorologists; these discuss precipitation, fog, icing, thunderstorms, and turbulence. The author is one of the nation's leading synoptic meteorologists, and the chapters on this subject are a gold mine of both classical and recent information and theories. This descriptive material and its presentation are alone quite enough to make *General Meteorology* an excellent book. The chapter on fog seemed especially good to the reviewer.

There are abundant references to the literature, and 300 well-drawn figures and maps illustrate the text. Three appendices give code tables, unit conversions, and a number of graphs useful to forecasters. The typography is excellent, and the reviewer found almost no misprints.

Professor Byers states in the preface that at present "More and more dynamic meteorology and synoptic meteorology lose their separate identities, and a general broad treatment of the subject becomes necessary." So we find that one-third (in general, the beginning) of the book is devoted primarily to such topics as atmospheric radiation balance, meteorological instruments and observations, thermodynamics, and the hydrodynamics of the atmosphere. The chapters on instruments are very good and as up-to-date as military censorship could permit in 1944. The dynamics material chosen for treatment is generally well-selected, and one topic (turbulence) is nicely blended with the related observational information. Some of the material, like Elsasser's radiation chart, is believed to appear here for the first time in a general textbook, and it is very welcome. One cannot feel, however, that the author has made much progress toward the truly difficult goal of integrating dynamic and synoptic meteorology in a "general broad treatment."

A more serious deficiency of the dynamics chapters results from the author's apparently limited skill or interest in writing mathematics and physics. In these chapters one finds too often that the definitions are loose, that the discussion is vapid, and that the mathematical derivations are clumsy. Worst of all, the physics is occasionally wrong (see, for example, the derivation of the Coriolis force at the top of page 192). While the inclusion of these chapters serves a very useful purpose, one can only wish that they were up to the standard of the rest of the book.

The book is especially recommended to persons interested in modern developments in descriptive and synoptic meteorology. It will also be well liked by many readers who want a general survey of meteorology complete in one volume.

A list of chapter titles follows: The Sun and the Earth; The Nature of Radiation; The Heat Balance of the Atmosphere; The Distribution of Temperature; Observations and Station Instruments; Upper-air and Cloud Observations; Thermodynamics and Statics; Horizontal Motion in the Atmosphere—the Winds; The General Circulation; Factors Determining Air Mass Structure; Air Masses of North America; Air Masses of Europe, Asia and the Southern Hemisphere; Frontogenesis and Frontal Structure; Fronts and Cyclones; Technique of Synoptic-chart Analysis; Highs, Lows and the Upper Air; Dynamic Effects in the Troposphere; Tropical Cyclones; Forecasting Technique; Condensation and Precipitation; Fog; The Formation of Ice on Aircraft; Thunderstorms and Related Phenomena; Atmospheric Turbulence and the Wind Structure Near the Surface of the Earth.

George E. Forsythe
Seattle, Washington

Elementary Statistics

By Hyman Levy and E. E. Preidel. Pp. 184, Figs. 29, 5¼×7¾ in. The Ronald Press Company, New York, 1945. Price $2.25.

This little book presents in simple and interesting fashion some of the fundamental concepts of statistics as taught by the authors at the Imperial College of Science, London. It appears to have been designed for students at about the sophomore level in a typical American engineering curriculum; the mathematics employed is for the most part algebra, although no attempt is made to avoid simple differentiation and integration where they can be used effectively.

The subjects covered are: types of frequency distributions; the arithmetic mean; measurement of dispersion; the accuracy of the mean; related measurements, correlation, and the correlation coefficient; elementary probability; the Gaussian "normal" distribution, and some of its properties; the binomial and Poisson distributions; and the elements of statistical quality control. The text is enlivened by many examples and a list of problems follows each of the fifteen chapters with answers at the back of the book.

The authors begin by making a distinction between a problem in which it is known that a true value exists although experimental difficulties prevent its determination, and one in which measurement presents no difficulty, but random variations in the quantity measured must be expected. As an example of each they cite "The current that will circulate in a wire of given size and material when connected to the two poles of a battery whose voltage difference is one volt," and "The number of visitors entering a museum per day." The practicability of such a differentiation is questionable. The existence of *a* true value of any physical property is debatable. In addition, it is doubtful that the system of causes of variability of current in an electric circuit is any less complex than that governing the number of museum visitors. The apparent difference in the two examples is caused by the great difference in "scale" of the units of measurement compared with the smallest integral units of the things measured. Unquestionably some data exhibit some of the characteristics of both classes. Thus it appears to the reviewer that this classification lacks a clear dividing line and is open to question, but it may be that, in questioning, the reader is led to an appreciation of the nature of data amenable to statistical treatment, which is, after all, the authors' purpose. Several minor typographical errors were found, particularly in exponents, but all are obvious mistakes.

Otherwise, this book should appeal to engineers and engineering students. Most American statistics texts fall into one of two classes: those designed for courses offered to sociology, economics, and education students, and those intended for mathematicians. The first have avoided all but the simplest algebra and have consequently been wordy with few developments or proofs. The others have been discouraging to the applied scientist, particularly if he is studying on his own. The book reviewed here is very readable and furnishes sufficient mathematical development to allow the student to use the results with confidence. It is recommended as an introduction to more advanced texts, and, as such, should contribute to the growing use of statistical methods.

D. K. Wright, Jr.
Case School of Applied Science

Engineering Mathematics

By Harry Sohon. D. Van Nostrand Company, Inc., New York, 1944.

The increasing use of analytical methods in engineering problems has been reflected in recent years in a rapid growth of textbook literature in applied mathematics. Among the contributions published during the war years is the book under review, a compact treatment of the mathematical background in algebra, differential equations and vector analysis, useful to the engineer. As the author states, this book "is intended to strengthen the student in algebra and to provide him with certain mathematical tools which depend on the calculus." The author is Assistant Professor of Electrical Engineering at the University of Pennsylvania.

Following a brief but clear presentation of interpolation formulae and techniques is a detailed treatment of determinants and matrices with many problems. Although the idea of a matrix and some of its properties are very clearly presented, it seems to the reviewer that an example or two indicating direct application to electric circuit theory or to a coupled vibrating system would prove highly instructive to the student engineer.

A chapter on dimensional analysis furnishes the student with an insight into a phase of applied mathematics which is altogether too often neglected in undergraduate engineering training. The numerous, complete illustrations of the methods give the reader an excellent idea of the power and utility of dimensional analysis.

Algebraic equations and their solution by exact and by approximate methods are treated in great detail in Chapters 5 and 6. It is unfortunate that Descartes' rule of signs for predicting the number of positive and negative roots of an algebraic equation is not mentioned. The inclusion of the ideas of continuity and derivative and the use of the latter in conjunction with the solution of algebraic equations lends strength to the presentation. The reviewer feels, however, that somewhat greater use of graphical aids would enhance the value of these chapters to the student.

For the most part, the discussion of ordinary differential equations and vector analysis is done in the usual manner. Gamma and Bessel's functions are introduced, and their properties are briefly discussed. Fourier series are discussed in Chapter 7, together with a numerical method of solving for the coefficients which seems to the reviewer somewhat inferior to other methods that have been devised. An interesting discussion of approximations to the Fourier coefficients concludes the chapter.

The book closes with a discussion of several problems in physics and engineering which illustrate admirably the use of the mathematical tools developed in the preceding chapters.

Taken as a whole, the book forms a compilation of mathematical background, which should enhance the powers of the serious student in his approach to the solution of engineering problems.

S. W. McCuskey
Case School of Applied Science

Infrared Spectroscopy

By R. B. Barnes, R. C. Gore, U. Liddel, V. Z. Williams. Pp. 236. Reinhold Publishing Corporation, New York. Price $2.25.

The early infra-red work by Coblentz in 1905, which pointed out the characteristic spectra of each molecule, laid the foundation for the present application of infra-red spectroscopy to the qualitative and quantitative analysis of chemical compounds, which is ably presented in the book *Infrared Spectroscopy*. This recent contribution from the American Cyanamid Research Laboratories is a worthy addition to the knowledge of infra-red spectroscopy because it points out the application of a relatively new physical tool in solving analytical problems in a large chemical industry. Unquestionably the book fulfills a long-felt need

for a readable text in this rapidly expanding field, which would appeal to chemists as well as physicists.

The material on pages 1 to 113 appeared first in *Industrial and Engineering Chemistry (Analytical Edition)*, Volume 15, No. 11, Pages 659 to 709 (1943) as an article by the present authors entitled "Infrared Spectroscopy, Industrial Applications." For formal publication a bibliography of 2701 references was added under an author and subject index on pages 114 to 236. This bibliography is a unique and highly useful section both to those now engaged in and about to enter the field of infra-red spectroscopy.

Included, also, is an extensive library of transmission curves on 363 different organic compounds with accompanying empirical formulae and alphabetical indices to these reference curves. Each of the spectral curves is mapped between 1000 and 2000 cm^{-1} in a space approximately 3 by 1.5 inches. While a few curves are given to 775 cm^{-1}, it is regretted that the frequency range was not extended to 3500 cm^{-1}. However, it is recognized that no prism material, such as LiF or CaF_2, was available for dispersion in this region at the time this work was done. Of special interest to physicists is the structural formula on the curve sheet of each compound.

The body of the book is divided into the following sections: (a) Qualitative Applications; (b) Quantitative Applications; (c) Experimental Equipment and Techniques; (d) Spectrometer; (e) Preparation of Samples; (f) Origin of Infrared Spectra; (g) Intensity of Absorption Bands; (h) Qualitative Analysis; (i) Quantitative Analytical Techniques; (j) Present Status of Infrared Spectroscopy in Industry. The theory of infra-red absorption and its relation to molecular structure are discussed to provide the essential background for detailed descriptions of techniques useful in analysis. Instead of presenting the mathematical theory of normal modes of vibration, the concept of normal modes in molecules is clarified by analogy with mechanical models. It helps those unfamiliar with infra-red spectroscopy to visualize the origin of the spectra.

A number of empirical correlations between bond frequencies and structural relationships are cited. For example, the characteristic C=O frequencies in different anhydrides, esters, acids, ketones and aldehydes, and the C=C frequencies in aliphatic and aromatic compounds are plotted for easy comparison. The quantitative analysis section, although not as detailed and specific as more recent publications, serves as an excellent introduction to this new infra-red application.

W. C. Sears
The Firestone Tire and Rubber Company

Climatology

By Bernard Haurwitz and James M. Austin. Pp. 410, Figs. 38. McGraw-Hill Book Company, Inc., New York. Price $4.50.

This book is intended to be an introduction to climatology for students of meteorology and it is assumed that the reader is familiar with the basic physical concepts of the subject. It is doubtful if it will prove of much interest to the average reader who has not had a previous introduction to meteorology since the description of the various climates is given in terms of the dynamics of the atmosphere, of air mass types, and of frontal activity.

The book is divided into two parts, the first of which presents a general survey of the various meteorological elements, their distribution over the globe, and their seasonal and diurnal variation. The factors that determine the distribution and variation of the meteorological elements are discussed.

The first chapter deals with insolation, its variation, its absorption by the atmosphere, and the heat balance over the earth maintained by radiation. In the second chapter, the temperature and its distribution over the surface of the earth as a function of latitude and time and the modifications introduced by the effects of the land and sea are discussed. Lapse rates and the temperatures at high altitudes are also briefly treated. A discussion of the winds and some of the important factors that modify them follows in the third chapter.

In the following chapters, precipitation as a climatological element is discussed at some length, also the precipitation as a function of the latitude, and the effect of various factors of time and space upon its distribution is explained. Evaporation and condensation are likewise treated. After a brief discussion of air masses, climatic zones and types of climates are outlined and then discussed in considerable detail. The first part ends with a chapter dealing with various microclimatology.

The second part of the book is devoted to regional climatology. The special climatic characteristics of each continent and of the oceans are set forth in considerable detail. The effects of the mountain chains, bodies of inland waters, and ocean currents on each of the continents are discussed. The air masses that affect each continent, their source regions, their characteristics, and modifying influences are given. The influence of the physical characteristics of the various regions of the earth is explained completely and clearly.

It will be a useful textbook for every meteorologist and student of weather and climate. It contains thirty-eight figures, most of them maps and charts; it has also eighty-three tables of illustrative and useful climatic information. A pocket inside the back cover contains seventeen plates of climatic information.

John G. Albright
Rhode Island State College

Frontiers in Chemistry. Volume III. Advances in Nuclear Chemistry and Theoretical Organic Chemistry

Edited by Robert Emmett and Oliver Grummitt. Pp. 165. Interscience Publishers, Inc., New York, 1945. Price $3.50.

This volume contains lectures delivered at Western Reserve University to an audience of graduate students and industrial chemists. The lecturers and their subjects are: A. S. Keston, "Isotopes and their Applications in

Biochemistry"; H. S. Taylor, "Applications of Isotopes in Catalytic Reactions at Surfaces"; H. R. Crane, "Techniques in Nuclear Physics"; L. G. S. Brooker, "Resonance and Organic Chemistry"; W. H. Rodebush, "The Hydrogen Bond and its Significance in Chemistry."

This book is expensive, considering its contents. The chapters by Taylor and Brooker represent excellent presentations of their subjects, although Brooker's chapter is somewhat deceptively titled since 46 of its 74 pages deal with the application of resonance theory to the problem of color and constitution. No doubt the other lectures were useful and interesting as lectures, but it seems an extravagance to present them in this expensive form since many more thorough textbook and reference treatments of the subjects are available.

A. L. ROBINSON
The University of Pittsburgh

Electron Optics and the Electron Microscope

BY V. K. ZWORYKIN, G. A. MORTON, E. G. RAMBERG, J. HILLIER, AND A. W. VANCE. Pp. 766+xi. Wiley and Sons, Inc., New York; Chapman and Hall, Ltd., London, 1945. Price $10.

This extensive and authoritative work has been compiled by the authors in recognition of a growing need for an instructive book on the theory, construction, and operation of the electron microscope. Their task has been made more difficult by the fact that a satisfactory presentation of electron optical theory and design requires a liberal use of mathematics that is frequently beyond the grasp of those who desire only to make intelligent use of the electron microscope as a research tool. In order not to discourage the practical electron microscopist, the authors have divided the book into two parts. The first is surprisingly non-mathematical while the second employs mathematics to a degree appropriate for methodical exposition.

Part I provides a qualitative introduction to electron optical principles and their application in numerous devices of such diverse character as the beam power tube, mass spectrograph, and cyclotron. Electron microscopes of all kinds are treated but the main emphasis is on the magnetic type which has proved the most successful. A chapter devoted to electron optics of high magnification contains much useful information for the practical electron microscopist. Discussion of lenses and their aberrations is undertaken in qualitative manner. Since special preparative techniques in electron microscopy are undergoing rapid development and improvement the description of this aspect of the subject lags, of necessity, behind present practice. A final chapter deals with applications and is profusely illustrated with electron micrographs representing research in many branches of science.

Part II contains a thorough, mathematical presentation of electron optical theory and design that is sufficiently broad to be of interest to many who are not primarily interested in electron microscopy. Chapters devoted to determination of potential distribution and trajectory tracing, and a section on ion optics, may be cited in this category. Electrostatic and magnetic lenses, electron mirrors, and their aberrations are treated systematically and at length. The discussion is supplemented by a consideration of the increase of electron mass in high velocity beams. Finally, present knowledge concerning image formation in the electron microscope is summarized in a sizable chapter embracing such topics as electron scattering in thin objects, diffraction effects, contour phenomena, and resolving power, all of which will be of interest from the practical as well as the theoretical point of view.

Except for the application of the electron microscope to research, which is to some extent beyond the scope of the book, the contents and bibliography are apparently complete to about the end of 1944. Liberal use is made of explanatory diagrams, charts, and micrographs throughout. The almost complete elimination of mathematical expression from the first part detracts from its value and the somewhat arbitrary division of the book tends to place obstacles in the way of the reader who attempts to correlate pertinent facts on some topics. In logical development, the two parts would be in the order indicated in the title. An historical note is injected occasionally but oddly enough the authors do not find it necessary, until page 493, to mention Busch, who first demonstrated the application of the lens formula in electron optics.

By gathering and presenting this material, some of it hitherto unpublished, the authors have performed a service to the rapidly expanding number of workers in the field. This book will be found useful to the beginner and the advanced worker alike and will undoubtedly stimulate progress in electron microscopy.

C. E. HALL
Massachusetts Institute of Technology

Seeing the Invisible

BY GESSNER G. HAWLEY. Pp. 195+xv, Figs. 71 (48 photos.), 13×19 cm. Alfred A. Knopf, New York, 1945. Price $2.50.

This little book is directed primarily towards those readers who have a sort of generalized curiosity about current advances in science, but whose interest is not sufficiently specific to allow them to be exposed profitably to a more precise and comprehensive discussion of the operation and use of the electron microscope. This much is already implied in the internally inconsistent title; only a scientifically uncritical mind would fail to realize that either the word "seeing" or the word "invisible" is used in an unusual and inaccurate sense, and that the purpose of the title is not to describe the book, but to catch the eye. However, the author himself in the preface leaves no doubt as to the purpose of the book, and consequently the technically informed reader should not be disappointed to find a considerable degree of semi-explanation and an almost overwhelming number of colorful adjectives.

If we look at the book then from the point of view of the reader to whom it is directed, we find that it presents certain attractive features. At the start it makes worthy effort to explain why we continue to be on the search for the ultimately small, and why this search has been

definitely limited in the field of optical microscopy. The discussion of resolving power, while somewhat vague owing to the avoidance of technical words commonly found in small dictionaries, is well done, and probably will explain for the first time to many uninitiated readers the importance of resolution in determining the useful magnifying power of optical instruments.

There follows a description of the construction and operation of the microscope, and then, greatly to the credit of Mr. Hawley, a chapter on the limitations of electron microscopy. Among the welter of popular books on the unlimited wonders of science, this reviewer finds as a most welcome novelty a chapter devoted exclusively to pointing out that there are still some hard nuts to crack in our progress toward knowing all! There are also several pages given to the history of the instrument, about which one gets the impression that the author has tried to be internationally minded, but has not been thoroughly informed as to the enormous priority of the developments of the German microscopists. For example, there is the statement that as late as 1944 one would have to say that the microscope was of "little or no use in determining the irregularities on the surface of metals." Anyone familiar with the early work of Mahl and others in Germany will know that surface replicas were successfully studied some years before 1944.

The expected descriptions of the use of the microscope in studying surface replicas, small particles, and biological objects comprise the last chapters. The fields covered are the ones most seriously studied to date, and the coverage is good for such a small book. The illustrations suffer badly, however, from being neither the most current nor the best ones available, and from being in many cases reprints of previously published illustrations. The "intelligent layman" should not be blamed for feeling, from an inspection of the illustrations, that the electron microscope mainly produces hazy, non-informative pictures on a gigantic scale of objects which we used to photograph hazily on a much smaller scale.

ROBLEY C. WILLIAMS
University of Michigan

Synthetic Rubber from Alcohol

BY A. TALALAY AND M. MAGAT. Pp. 298. Interscience Publishers, Inc., New York. Price $5.00

This treatise deals with the Russian method of making synthetic rubber from alcohol and is based principally on a survey of the Russian literature on this timely subject. Theoretical aspects, technology, and production practice on the conversion of alcohol to butadiene and the subsequent polymerization to polybutadiene are given, which makes this a valuable book for research chemists, physicists, and chemical engineers who are working in these fields. The book is of value for these reasons in spite of its limited field which does not cover the two-step process of converting alcohol to butadiene, which is used in the United States, nor does it cover compounding of the finished polymers.

The first chapter contains a discussion of the chemistry and theoretical mechanisms involved in the Lebedev Process for the direct conversion of alcohol to butadiene by the use of a dual catalyst producing simultaneous dehydration and dehydrogenation of ethanol resulting in a commercially satisfactory yield of butadiene. Considerable theory is given regarding the process utilizing the Lebedev concepts of "*a*" and "*b*" fragments and how they combine to produce not only butadiene but the by-products which are formed in the process. Gorin's work, which led to the recirculation of aldehyde with rectified alcohol to substantially increase the yields of butadiene, is mentioned. Methods of purification, temperature effects, as well as the use of catalysts and their theory of action, are given in some detail.

The second chapter concerns the technology of the S.K. Process. This chapter is particularly well written as it develops the process from the laboratory stage straight through the pilot plant stage to the final commercial production installation. Sufficient diagrams and flow sheets are given along with the description to make this a real contribution. Temperature, catalysts, yields, by-products, purification, and methods of analytical control, as well as explosion hazards, are covered.

The third and fourth chapters of the book deal primarily with the polymerization of the butadiene to polybutadiene. The theory of the processes leading to the production of macromolecules, such as polycondensation and polymerization, are explained in some detail along the concepts of "ionic" and "free radical" type polymerization catalysts. Autopolymerization, along with polymerization in liquid and vapor phases, is gone into in some detail. Emulsion polymerization is also described. A complete description of the "rod" and "rodless" methods of sodium polymerization are given along with the yields and effects of impurities.

The fourth chapter is concerned mainly with the physiochemical properties of the polymer. Considerable information on molecular weight, and molecular weight distribution, is given along with the various methods of determining molecular weight. The study of the properties of solutions and cements includes not only a study of viscosity, molar weight, and structure, but also a discussion of adhesion of these cements.

All of the chapters of this book are written in a clear and concise manner. A resumé of the contents of the chapter is given in the first paragraph, followed by a detailed description of the subject matter. Sufficient details are given in most cases so that laboratory set-ups can be duplicated, and sufficient flow sheets are given so that plant operations can be followed through. This book will be of particular interest to research chemists and physicists working in high polymer and related fields, not only because of the valuable information and logical presentation, but because of the large number of literature citations of which there are some five hundred and eighty in number. Complete author and subject indexes are included.

R. S. HAVENHILL
St. Joseph Lead Company

Dr. W. C. Röntgen

By Otto Glasser. Pp. 169, Illustrated, 21×14 cm. Charles C. Thomas, Springfield, Illinois, 1945. Price $4.50.

This little volume, sponsored jointly by the American Roentgen Ray Society, the Radiological Society of North America, and the American College of Radiology, has been prepared to celebrate the one-hundredth anniversary of the birth of W. C. Röntgen, and the fiftieth of his discovery of x-rays. It gives a brief, well-illustrated account of his life and work. A considerable portion of the text is occupied by new translations of the three classic papers in which Röntgen first announced the discovery of the rays and some of his subsequent investigations of their properties. They read a little more smoothly than those made by G. F. Barker in 1896 and 1897 which can be found reprinted in part in the *American Journal of Physics* **13**, 281 (1945), but a quick comparison discloses no essential difference in meaning.

The whole book is based largely on Glasser's earlier volume, *Wilhelm Conrad Röntgen and the Early History of the Roentgen Rays*, which is detailed and authoritative. Some new information is included, derived from "numerous letters, documents, and other communications obtained subsequent to that publication." A reader who is not making an intensive study of the subject, however, will gain from the smaller volume a clearer and more human picture of Röntgen, and will appreciate better the brilliance of the mind which led him, step by step, in inevitable logical sequence, through the series of experiments by which he convinced himself of the existence of a new kind of rays.

His first three papers on x-rays are models of clarity and straightforward exposition. Seldom was he led aside from the main theme by mistaking as significant an observation which today would be classed as trivial; he could seize upon an observation, select it like a magician from the worthless superficialities which threatened to conceal its true value, and devise crucial experiments from which a new property of the rays would emerge with such directness that the process of scientific discovery, to the uninitiated, was made to appear one of the simplest of man's activities.

Perhaps it was this masterly but deceptive simplicity of his experiments and writings which caused some less able minds to be skeptical, to be critical, and to belittle his achievements, and made Röntgen, at the same time, oversensitive to these adverse buffets of fortune. Perhaps, too, some of the misunderstandings arose from the haste with which his discovery, after its first publication, was transmitted to the whole civilized world, often by individuals who could neither understand nor evaluate it. In an apparently reputable book, published in 1896, one finds, for example, Röntgen's papers catalogued under the heading *X-rays and a New Art*. Whether or not the compiler intended this as a translation of the original title, *Eine Neue Art von Strahlen*, some seeds of misinterpretation were undoubtedly sown.

In reading about Röntgen's life, scientists will find that some difficulties of their profession have not changed or vanished in the intervening fifty years. They will understand his high hopes in moving to a new position, and his disappointment if, as sometimes happened, the equipment did not come up to expectations. They will be interested in his struggle to gain a foothold on the lowest rung of the academic ladder, an achievement that was made difficult by his somewhat unorthodox educational background. They will recognize a modern parallel in the fact that ability can very often be spelled without the letters Ph.D.

The book is strongly recommended, especially to young scientists.

Thomas H. Osgood
Michigan State College

New Booklets

The *General Radio Experimenter* for February, 1946, features "The Constant Waveform Frequency Meter." Eight pages. Published by General Radio Company, 275 Massachusetts Avenue, Boston 39, Massachusetts.

The Chemical Publishing Company, Inc., 26 Court Street, Brooklyn 2, New York, has issued a new catalog of technical books of all kinds. The catalog gives the date of publication of each book as well as price, number of pages, detailed descriptions, and full table of contents. Free on request.

Wabash Corporation, Photolamp Division, 335 Carroll Street, Brooklyn 31, New York, has published the latest phototechnical data on the complete Wabash Photolamp line with detailed exposure information, in Bulletin No. 743P. Photoflash and photoflood tables for use with all black and white and color films are included, with a special section devoted to color photography with both standard and daylight-blue photoflash bulbs. Four pages. Free on request.

Miniature Precision Bearings of Keene, New Hampshire, manufacturers of miniature radial, pivot, and special bearings, precision parts and tools, have issued a six-page leaflet entitled "Ball Bearings for Precision Instruments." Free on request.

Bulletin of the Atomic Scientists of Chicago for February 15 discussed "Operation Crossroads," and contained the fourth in a series of articles on atomic energy controls: "Aerial Survey—An Aid in Mining Control." The office of this organization is in Room 6, Social Science Building, University of Chicago, 1126 East 59th Street, Chicago, Illinois.

North American Philips Company, Inc., 100 East 42 Street, New York 17, New York, has recently published a 16-page booklet entitled *Phosphors and Their Behavior in Television*. Photographs and graphs included.